

Induction Level Determines Signature of Gene Expression Noise in Cellular Systems

Julia Rausenberger*

Institute of Physics and Faculty of Biology, University of Freiburg, Germany

Christian Fleck, Jens Timmer

Institute of Physics, University of Freiburg, Germany

Markus Kollmann

Institute for Theoretical Biology, Humboldt-University Berlin, Germany

Abstract

Noise in gene expression, either due to inherent stochasticity or to varying inter- and intracellular environment, can generate significant cell-to-cell variability of protein levels in clonal populations. We present a theoretical framework, based on stochastic processes, to quantify the different sources of gene expression noise taking cell division explicitly into account. Analytical, time-dependent solutions for the noise contributions arising from the major steps involved in protein synthesis are derived. The analysis shows that the induction level of the activator or transcription factor is crucial for the characteristic signature of the dominant source of gene expression noise and thus bridges the gap between seemingly contradictory experimental results. Furthermore, on the basis of experimentally measured cell distributions, our simulations suggest that transcription factor binding and promoter activation can be modelled independently of each other with sufficient accuracy.

*To whom the correspondence should be addressed. E-mail: juhe@fdm.uni-freiburg.de

Introduction

Within a genetically identical population, individual cells show significant phenotypic heterogeneity (1; 2; 3). This variability directly affects the cell's ability to respond to environmental factors like changes in ligand concentration. Especially, reactions underlying protein synthesis are often based on small numbers of molecules, like transcription factors or ribosomes, such that stochastic fluctuations have to be taken into account.

A lot of effort has been undertaken to quantify the origins of gene expression noise experimentally and theoretically. Stochasticity or noise inherent to gene expression seems to be one of the main driving forces for the observed cell-to-cell variability in several experiments which have measured the variance in protein abundances in different cellular systems (4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15; 16; 18). Considerable confusion stems from diverging experimental results which have identified different origins for the main contribution to gene expression noise (19) such that a complete picture is still missing. For prokaryotes, translational efficiency was identified as the main source of variability of expression levels consistent with a stochastic model in which proteins are produced in sharp and random bursts (20). However, later experimental observations in individual living cells either by measuring mRNA levels or by real-time observations at single molecule level indicated that promoter activation predominantly causes gene expression noise (21; 22). Furthermore, extrinsic factors, like the cellular state, were also identified to give the main contribution to phenotypic variations in a clonal population (16). Similar contradictory results have been found in eukaryotes, where in the budding yeast *Saccharomyces cerevisiae* a two-reporter system, expressing two fluorescent proteins from identical promoters, identified switching between active and inactive promoter states due to slow stochastic chromatin-remodelling events as the by far largest source of noise (5). In later *in-vivo* experiments it was shown for a large set of genes at their native expression levels that the noise has a clear sign of transcriptional origin due to low-copy mRNA molecules (7; 14). Moreover, a direct monitoring of mRNA production from a gene at the resolution of single molecules in mammals revealed strong mRNA bursts dominating gene expression noise (9). In contrast, for human cells, genes at native induction level showed significant noise contribution from long-term variations of the cellular state (15). It seems on first sight that no general rule can be given to determine the main sources of gene expression noise. Protein levels, however, are

often strongly optimized, because they have to allow for precise and reliable information processing within the cell. Any significant deviation from the optimal level would result in reduction of fitness and an evolutionary disadvantage. Thus, random fluctuations are in general detrimental for cellular systems and several regulatory mechanisms have evolved to minimize them. Only in rare cases noise can be used to drive phenotypic switching providing a non-genetic mechanism to population heterogeneity, as found for bacterial persistence against antibiotics (23) and competence for DNA uptake from the environment (24).

In order to track down the individual contributions of the molecular mechanisms involved in protein synthesis several mathematical models have been introduced (25; 26; 27; 28; 29; 30; 31). Some of these models ignore the effect of binomial partitioning by cell division which will lead to strong discrepancies to experiments whenever cellular mRNA is long-lived and appears in low copy number (29; 30; 31), whereas others lack the dynamic description of mRNA bursts (25; 27). In the present work we develop an analytical framework which allows for a time-dependent description of gene expression and accounts for effects of symmetric cell division. We consider a *one-gene-system* consisting of activator/transcription factor (TF) binding (repressor unbinding), promoter activation, transcription, and translation (Fig. 1). All gene specific events contribute to the so-called *intrinsic noise*. Differences between cells, either in global cellular state or in the concentration or activity of any factor that affects gene expression are referred to as *extrinsic noise* (4). Therefore, the cell-to-cell variability of a specific protein in a large clonal population with fixed generation time is characterized by the two contributions of intrinsic and extrinsic noise, summing up to the overall variance of the protein. Assuming no specific feedback of a produced protein on upstream processes, the intrinsic noise contribution decomposes into partial contributions stemming from activator binding, promoter activation, transcription and translation. In deriving analytical expressions for these partial contributions to gene expression noise, we discuss limiting cases for mRNA and protein lifetimes. We show that the magnitudes of the different noise contributions depend strongly on induction level, synthesis rates, and molecule lifetimes associated with each individual gene. Therefore, differences in the induction level, e.g., due to different experimental set-ups, might provide a possible explanation for the diverging experimental findings of cell-to-cell variations even in the same organisms. Furthermore, on the basis of experimentally measured cell distributions of wild-type and over-expressed cells of *E. coli*, our simulations propose that, in prokaryotes, activator binding and promoter activation are

independent of each other and thus can be modelled for a good approximation separately.

Results

Minimal model for gene expression noise

Four major steps are involved in a generic model of gene expression in living cells (30): (i) activator/TF binding (repressor unbinding), (ii) promoter (DNA) activation, (iii) transcription, and (iv) translation. In the present work we follow a previous approach (30) and model synthesis and degradation of mRNA and protein by a *birth-and-death* process. Activator binding and promoter activation are described as a random telegraph process, because they are assumed to switch randomly between zero and one with exponentially distributed waiting times (9; 21). The state of the activator is given by the stochastic variable $B(t)$ switching between $B(t) = 1$ and $B(t) = 0$, if the activator is bound or unbound, respectively. Promoter activation can be expected to occur on much slower time scales than activator binding (21; 22; 32; 33) such that the time scales can be separated. Therefore, we approximate activator binding, $B(t)$, as an equilibrated binding process. The Master equation for the probability to find a promoter in its active, $P(1, t)$, or inactive, $P(0, t)$, state is given by

$$\partial_t P(1, t) = -\lambda_A^- P(1, t) + \lambda_A^+ P(0, t), \quad (1)$$

where λ_A^+ represents the switching rate from the inactive to the active state and λ_A^- the rate for the inverse process. We define

$$\gamma_A = \lambda_A^+ + \lambda_A^- \quad (2)$$

and denote the initial state of the promoter at time t_0 by α_0 . The solution for (1) with initial conditions $P(1, t_0 | \alpha_0, t_0) = \delta_{1, \alpha_0}$ reads:

$$P(1, t | \alpha_0, t_0) = \frac{\lambda_A^+}{\gamma_A} + \left(\delta_{1, \alpha_0} - \frac{\lambda_A^+}{\gamma_A} \right) e^{-\gamma_A(t-t_0)}. \quad (3)$$

We can assign the stochastic variable $A(t) \in \{0, 1\}$ to describe the state of the promoter. The process of promoter activation seems to have no significant correlations with the cell cycle (9; 21). Therefore, we can employ a stationary solution for the auto-covariance function

$$\langle A(t)A(t') \rangle - \langle A(t) \rangle \langle A(t') \rangle = \text{var}(A) e^{-\gamma_A |t-t'|} \quad (4)$$

with $\text{var}(A) = \lambda_A^+ \lambda_A^- / \gamma_A^2$, and $\langle A \rangle = \lambda_A^+ / \gamma_A$. The mean $\langle A \rangle$ can be interpreted as a measure of the fraction a promoter spends in its active state for which holds $0 \leq \langle A \rangle \leq 1$.

Transcription and translation are modelled as a *birth-and-death* process. The corresponding mRNA and protein trajectories are denoted by $R(t)$ and $X(t)$, respectively. For the conditional probability $P(n_i, t | n'_i, t_0)$ to observe n_i molecules ($n_1 = R, n_2 = X$) in a given cell at time t we obtain the stochastic Master equation

$$\begin{aligned} \partial_t P(n_i, t | n'_i, t_0) &= \lambda_i^+(t) P(n_i - 1, t | n'_i, t_0) \\ &+ (n_i + 1) \lambda_i^- P(n_i + 1, t | n'_i, t_0) \\ &- (\lambda_i^- n_i + \lambda_i^+(t)) P(n_i, t | n'_i, t_0) \end{aligned} \quad (5)$$

where n'_i molecules are observed initially at time t_0 . Synthesis rates are given by $\lambda_1^+(t) = \lambda_R^+ A(t) B(t)$ for the mRNA or $\lambda_2^+(t) = \lambda_X^+ R(t)$ for the proteins. The corresponding degradation rates, $\lambda_1^- = \lambda_R^-$ and $\lambda_2^- = \lambda_X^-$, are assumed to be independent of molecule number and constant in time. Furthermore, the trajectories $A(t)$ and $B(t)$ are assumed to be independent of each other which will be justified below. Transcription at time t is only possible if the activator is bound *and* the promoter is in the active state, $A(t)B(t) = 1$, cf. Fig. 1. We assume that the translation rate of specific proteins depends linearly on the actual amount of the corresponding mRNA, $R(t)$.

A lot of effort has been undertaken to generate time resolved expression data of single cells (15). The mathematical description of gene expression becomes more complicated if cell division is included, because cell division introduces another important time scale into the system, the generation time T_G . Previous theoretical approaches avoided this problem by assuming an increased protein degradation rate instead of explicitly taking cell division into account (29). Such an approximation might be justified for systems where effects of binomial partitioning can be neglected and cellular growth and protein synthesis scale linearly in time. In this case the concentrations - but not the absolute copy number of molecules - are approximately constant over cell cycles. However, one should note that experiments measure the total molecule number instead of concentrations. Measurement of concentrations would require precise knowledge of the individual cell volume at any time. Recent experiments show that the main effect of cell division seems to be binomial partitioning of molecules (16; 21) such that the assumption of an increased protein degradation does not hold.

If we assume symmetric cell division at time t_0 with $t' < t_0 < t$ and a binomial distribution

of the molecules, the following relationship for the conditional probability $P(n_i, t|n'_i, t')$ holds:

$$P(n_i, t|n'_i, t') = \sum_{n_i^-=0}^{\infty} \sum_{n_i^+=0}^{n_i^-} P(n_i, t|n_i^+, t_0) \binom{n_i^-}{n_i^+} 2^{-n_i^-} P(n_i^-, t_0|n'_i, t'). \quad (6)$$

The amounts of molecules before and after the last cell division occurring at time t_0 are represented by n_i^- and n_i^+ , respectively. To solve Eqn. (5) we use the generating function $G(s, t|n'_i, t') = \sum_{n_i=0}^{\infty} s^{n_i} P(n_i, t|n'_i, t')$ and assume a finite number of cell divisions between t' and t . After some algebra we arrive (for derivation see SI) at:

$$G(s, t|n'_i, t') = \left[2^{-D(t, t')} (s - 1) e^{-\lambda_i^-(t-t')} + 1 \right]^{n'_i} \times \exp \left[(s - 1) \int_{t'}^t \lambda_i^+(t'') e^{-\lambda_i^-(t-t'')} 2^{-D(t, t'')} dt'' \right], \quad (7)$$

where $D(t, t')$ is the number of cell divisions that have occurred between two time points $t > t'$. The mean amount of molecules at time $t > t'$ can be calculated via $\partial_s G|_{s=1}$. If we assume that initially the process is Poissonian-distributed the mean amount of molecules is given by

$$\langle n_i(t) \rangle_i = \langle n_i(t') \rangle_i 2^{-D(t, t')} e^{-\lambda_i^-(t-t')} + \int_{t'}^t \lambda_i^+(t'') e^{-\lambda_i^-(t-t'')} 2^{-D(t, t'')} dt'' \quad (8)$$

where the average over R or X is denoted by $\langle \cdot \rangle_i$. We define the generating function $\Theta(s, t|n'_i, t_0) = \sum_{n_i=0}^{\infty} s^{n_i} P(n_i, t|n'_i, t_0)$ and use Eqn. (7) to obtain the auto-correlation function ((17)) via $\partial_s \Theta|_{s=1}$ (for detailed derivation see SI):

$$\langle n_i(t), n_i(t') \rangle_i = \langle n_i(t') \rangle_i 2^{-D(t, t')} e^{-\lambda_i^-(t-t')}. \quad (9)$$

Eqn. (9) shows that each cell division halves the magnitude of the auto-correlation function. Replacing n_i by R or X , we obtain the mean amount and the auto-correlation function of the mRNA or the protein from Eqns. (8) and (9), respectively.

Partial contributions to gene expression noise

The heterogeneity in gene expression of a population can be quantified using the standard deviation (σ) divided by the mean (μ), i.e., $\eta = \sigma/\mu$. The quantity η is commonly denoted as noise and provides a physiologically relevant measure of gene expression variability as it quantifies relative fluctuations independent of the expression level. Two main contributions

to the overall variance, σ_{tot}^2 , determine the cell-to-cell variability in the amount of a protein, X : the intrinsic variance, σ_I^2 , which is distinctive for each gene in its genomic context and an extrinsic noise contribution stemming from the variance σ_E^2 , which is due to fluctuations in the intra-cellular environment. If we do not assume any significant feedback of the expressed protein on extrinsic factors (cf. Fig. 1 and (34)), the overall variance in a large clonal population of cells with fixed generation time T_G sums up to:

$$\sigma_{tot}^2 = \sigma_I^2 + \sigma_E^2. \quad (10)$$

The extrinsic contributions are usually separated into fluctuations of upstream factors that drive expression directly, e.g., a given activator concentration and the cellular state that influences gene expression, e.g., via variations in polymerase, ribosome, and protease concentrations. The average over all possible trajectories of protein copy number $X(t)$, mRNA copy number $R(t)$, promoter activation $A(t)$, and activator binding $B(t)$ is defined by $\langle \cdot \rangle_{\mathbf{I}} = \int \cdot P[X(t); R(t); A(t); B(t)] dX(t) dR(t) dA(t) dB(t)$ with the multi-index $\mathbf{I} = (X, R, A, B)$, using e.g., $\langle \cdot \rangle_X = \int \cdot P[X(t)|R(t)] dX(t)$. Note that $P[X; R; A; B]$ can be written as $P[X|R]P[R|A, B]P[A]P[B]$. We can identify different contributions to the *intrinsic noise* of protein synthesis:

$$\begin{aligned} \sigma_I^2 &= \langle (X(t) - \langle X(t) \rangle_{X,R,A,B})^2 \rangle_{X,R,A,B} \\ &= \langle (X(t) - \langle X(t) \rangle_X)^2 \rangle_{X,R,A,B} \\ &\quad + \langle ((\langle X(t) \rangle_X - \langle X(t) \rangle_{X,R})^2) \rangle_{R,A,B} \\ &\quad + \langle ((\langle X(t) \rangle_{X,R} - \langle X(t) \rangle_{X,R,A})^2) \rangle_{A,B} \\ &\quad + \langle ((\langle X(t) \rangle_{X,R,A} - \langle X(t) \rangle_{X,R,A,B})^2) \rangle_B \\ &= \sigma_B^2 + \sigma_A^2 + \sigma_R^2 + \sigma_X^2, \end{aligned} \quad (11)$$

$$= \sigma_B^2 + \sigma_A^2 + \sigma_R^2 + \sigma_X^2, \quad (12)$$

where the right-hand-side denotes the sum over variances (for derivation see SI) corresponding to the processes of activator binding (σ_B^2), promoter activation (σ_A^2), transcription (σ_R^2), and translation (σ_X^2), respectively. The summing up of the individual variances is only possible if there are not any feedbacks from downstream to upstream processes. We find for the noise contribution due to translation: $\eta_X^2 = \sigma_X^2 / \langle X \rangle_{\mathbf{I}}^2 = 1 / \langle X \rangle_{\mathbf{I}}$. Also, the binomial distribution of the proteins caused by cell division converges quite rapidly to a Gaussian distribution for an increased amount of molecules. The amount of protein synthesized

per mRNA can be estimated to be of the order $10 - 10^3$ for most systems. Thus, for the experiments considered in this work, we can neglect translational noise in comparison to other noise contributions.

In order to make the analytical calculations feasible, we assume a fixed induction level of the activator, $B(t) = B_{eq}$. Due to the separation of the time scales (fast activator binding and very slow promoter activation), fluctuations from activator binding can be neglected compared to fluctuations from promoter activation and transcription such that the averaging over B cancels out in Eqns. (11). Therefore, noise contribution from activator binding, σ_B^2 , is not present and we set $\mathbf{I} = (X, R, A)$. Explicit expressions for the different intrinsic noise contributions of promoter activation, σ_A^2 , and transcription, σ_R^2 , can be obtained from Eqns. (4),(8) and (9) (see SI for a detailed derivation). We introduce the abbreviations

$$\mathcal{A} = \langle A \rangle B_{eq} \lambda_R^+ \lambda_X^+ \quad (13)$$

representing the average acceleration of protein synthesis in absence of any degradation, and

$$Z_R(T_G) = \frac{1 - \exp[-\lambda_R^- T_G]}{1 - 2^{-1} \exp[-\lambda_R^- T_G]}. \quad (14)$$

Assuming infinitely many cell divisions and averaging over all intrinsic processes yields for the mean amount of protein and mRNA:

$$\begin{aligned} \langle X(t - t_0) \rangle_{\mathbf{I}} = & \frac{\mathcal{A}}{\lambda_R^-} \left[\frac{1}{\lambda_X^-} \left(1 - e^{-\lambda_X^- (t-t_0)} \right) - \frac{1 - 2^{-1} Z_R}{\lambda_R^- - \lambda_X^-} \left(e^{-\lambda_X^- (t-t_0)} - e^{-\lambda_R^- (t-t_0)} \right) \right] \\ & + \frac{2^{-1}}{1 - 2^{-1} e^{-\lambda_X^- T_G}} e^{-\lambda_X^- (t-t_0)} \frac{\mathcal{A}}{\lambda_R^-} \left[\frac{1}{\lambda_X^-} \left(1 - e^{-\lambda_X^- T_G} \right) - \frac{1 - 2^{-1} Z_R}{\lambda_R^- - \lambda_X^-} \left(e^{-\lambda_X^- T_G} - e^{-\lambda_R^- T_G} \right) \right] \end{aligned} \quad (15)$$

$$\langle R(t - t_0) \rangle_{\mathbf{I}} = \langle A \rangle B_{eq} \frac{\lambda_R^+}{\lambda_R^-} \left[1 - (1 - 2^{-1} Z_R) e^{-\lambda_R^- (t-t_0)} \right]. \quad (16)$$

The terms involving the generation time, T_G , reflect the memory of the contributions that have been generated in previous generations and have been passed over to the actual generation. Since we account for cell division, the results can be directly compared to time-resolved expression data of protein levels.

In the following we discuss three important physiological limiting cases in detail. The derived asymptotic expressions for the mean and the variances are valid for any time t within a given cell cycle, $0 \leq t - t_0 \leq T_G$, where t_0 denotes the time point of the last cell

division. In the first two cases we omit the explicit formulas for the corresponding noise contributions $\eta = \sigma/\mu$, since these formulas are easily derived but do not contain new information with respect to the derived mean and variances.

Case I: Short mRNA lifetimes and long protein lifetimes, $(\lambda_R^-)^{-1} \ll T_G \ll (\lambda_X^-)^{-1}$. This is the most likely physiological case. We obtain for $t - t_0, T_G \gg (\lambda_R^-)^{-1}$ the following asymptotic expressions for the average amount of protein and mRNA:

$$\langle X(t - t_0) \rangle_{\mathbf{I}} = \frac{\mathcal{A}}{\lambda_R^-} [(t - t_0) + T_G] \quad (17)$$

$$\langle R(t - t_0) \rangle_{\mathbf{I}} = \langle A \rangle B_{eq} \frac{\lambda_R^+}{\lambda_R^-}. \quad (18)$$

Note that Eqn. (17) implies a linear increase of the mean amount of protein in time as well as a doubling of protein synthesized over one cell cycle, $\langle X(T_G) \rangle_{\mathbf{I}} = 2\langle X(t_0) \rangle_{\mathbf{I}}$. The stationary mRNA level is recovered immediately after cell division. The noise contributions from transcription and promoter activation are also time-dependent (for derivation see SI) and read in the limit $t - t_0, T_G \gg (\lambda_R^-)^{-1}, \gamma_A^{-1}$:

$$\sigma_R^2(t - t_0) = 2\mathcal{A} \frac{\lambda_X^+}{(\lambda_R^-)^2} \left[t - t_0 + \frac{1}{3}T_G \right] \quad (19)$$

$$\sigma_A^2(t - t_0) = 2\mathcal{A}^2 \frac{\text{var}(A)}{\langle A \rangle^2} \frac{1}{(\lambda_R^-)^2 \gamma_A} \left[(t - t_0) + \frac{1}{3}T_G \right]. \quad (20)$$

Case II: Long mRNA lifetimes and long protein lifetimes, $(\lambda_R^-)^{-1}, (\lambda_X^-)^{-1} \gg T_G$. Molecule lifetimes are significantly larger than the generation time and the switching rate of promoter activation, which result in the asymptotic expressions for the means:

$$\langle X(t - t_0) \rangle_{\mathbf{I}} = \frac{1}{2}\mathcal{A} [(t - t_0)^2 + T_G^2] \quad (21)$$

$$\langle R(t - t_0) \rangle_{\mathbf{I}} = \langle A \rangle B_{eq} \lambda_R^+ [t - t_0 + T_G]. \quad (22)$$

The mean amount of mRNA increases linearly whereas the mean amount of protein increases quadratically within one cell cycle. Both the mRNA and the protein synthesized over one cell cycle are doubled. In this asymptotic case, dilution due to cell division is the only mechanism which determines the correlation times. Regarding transcriptional noise, we find

in the limit $t - t_0, T_G \gg \gamma_A^{-1}$ strong contributions from previous generations, reflected by several terms involving T_G :

$$\sigma_R^2(t - t_0) = \frac{1}{3} \mathcal{A} \lambda_X^+ \left[(t - t_0)^3 + T_G^2(t - t_0) + \frac{2}{3} T_G^3 \right]. \quad (23)$$

Noise due to promoter activation in the limit $t - t_0, T_G \gg \gamma_A^{-1}$ takes the functional form:

$$\sigma_A^2(t - t_0) = \frac{2}{3} \mathcal{A}^2 \frac{\text{var}(A)}{\langle A \rangle^2} \frac{1}{\gamma_A} \left[(t - t_0)^3 + T_G(t - t_0)^2 + 12 T_G^2(t - t_0) + 14 T_G^3 \right]. \quad (24)$$

Experimentally one could test this case using a set-up with artificially stabilized mRNA.

Case III: Short mRNA lifetimes and short protein lifetimes, $(\lambda_R^-)^{-1}, (\lambda_X^-)^{-1} \ll T_G$. In the exceptional case that both protein and mRNA lifetimes are significantly shorter than the generation time, memory over generations is eliminated and the stationary protein level is recovered immediately after cell division. The solutions are consequently time-independent and agree with those found earlier by Paulsson (29; 30) for the mean amount of protein and mRNA:

$$\langle X \rangle_{\mathbf{I}} = \frac{\mathcal{A}}{\lambda_R^- \lambda_X^-} \quad (25)$$

$$\langle R \rangle_{\mathbf{I}} = \langle A \rangle B_{eq} \frac{\lambda_R^+}{\lambda_R^-} \quad (26)$$

and the variances

$$\sigma_R^2 = \mathcal{A} \frac{\lambda_X^+}{\lambda_X^- \lambda_R^- (\lambda_R^- + \lambda_X^-)} \quad (27)$$

$$\sigma_A^2 = \frac{\mathcal{A}^2 \text{var}(A)}{\langle A \rangle^2 \lambda_R^- \lambda_X^- (\gamma_A + \lambda_R^-) (\lambda_R^- + \lambda_X^-)} \left(1 + \frac{\lambda_R^-}{\gamma_A + \lambda_X^-} \right). \quad (28)$$

The noise contributions from transcription and promoter activation to gene expression noise are given by

$$\eta_R^2 = \frac{1}{\langle R \rangle_{\mathbf{I}}} \frac{\lambda_X^-}{\lambda_R^- + \lambda_X^-} \quad (29)$$

$$\eta_A^2 = \frac{\text{var}(A)}{\langle A \rangle^2} \frac{\lambda_R^-}{\gamma_A + \lambda_R^-} \frac{\lambda_X^-}{\lambda_R^- + \lambda_X^-} \left(1 + \frac{\lambda_R^-}{\gamma_A + \lambda_X^-} \right) \quad (30)$$

and have been calculated by Paulsson as the normalized stationary variance with the same result ((30), Eqn. (4)).

Noise regimes account for different experimental observations

Different origins of noise have been proposed and measured by several experimental groups (5; 7; 9; 14; 15; 16; 20; 21; 22). Recently, Kaufmann and van Oudenaarden critically reviewed these experimental observations (19). The contradictory results, even in the same eukaryotic organism *S. cerevisiae*, support the idea that gene expression is influenced by more than one main driving source. In the following we focus on the budding yeast *S. cerevisiae* and the contradictory experimental results found by Bar-Even et al., Newman et al. and Raser and O’Shea (5; 7; 14). Bar-Even et al. as well as Raser and O’Shea developed the same mathematical model to describe gene expression: it contains the processes of gene/promoter activation, transcription and translation (see Supporting Online Materials of (5) and (7)). However, the authors interpreted the theoretical results according to their experimental observations. Raser and O’Shea (5) measured the intrinsic noise strength of the *PHO5* and *PHO84* promoters at different rates of gene expression in promoter constructs. They concluded that the noise intrinsic to gene expression is promoter-specific: noise generation at the *PHO5* promoter depends on stochastic promoter activation due to chromatin remodelling. In contrast, Bar-Even et al. (7) investigated native expression of 43 genes in each of 11 conditions, whereas Newman et al. (14) presented an extensive overview of protein noise for more than 2500 proteins expressed from their endogenous promoter and natural chromosomal position by the use of a combination of high-throughput flow cytometry and a library of GFP-tagged yeast strains. Both latter studies concluded that random *birth-and-death* of low-copy mRNA molecules describe the large observed variations quite well: for the great majority of proteins the noise level is inversely proportional to the mean protein abundance implying a clear signature of a Poisson process. The obvious question arises: what is the predominant source of noise: promoter activation, as suggested by Raser and O’Shea, or mRNA fluctuations due to low copy number as proposed by Bar-Even et al. and Newman et al.? An explanation for these contradictory experimental results is given by our stochastic model considering activator binding explicitly. For approximately constant activator concentration and high amount of protein synthesized per mRNA, the intrinsic variance σ_I^2 of Eqn. (12) reduces to $\sigma_I^2 = \sigma_A^2 + \sigma_R^2$. Hence the ratio σ_A^2/σ_R^2 determines the predominant source of noise: if $\sigma_A^2/\sigma_R^2 \gg 1$, promoter activation will be the dominant process while in the case of $\sigma_A^2/\sigma_R^2 \ll 1$ the major part of gene expression noise is due to

transcription. For long protein lifetimes and short mRNA lifetimes (Eqns. (19), (20)) the ratio is given by

$$\frac{\sigma_A^2}{\sigma_R^2} = B_{eq} \frac{\text{var}(A)}{\langle A \rangle} \frac{\lambda_R^+}{\gamma_A} = B_{eq} \frac{\lambda_A^- \lambda_R^+}{\gamma_A^2}. \quad (31)$$

Note that this ratio is time- and cell cycle-independent although σ_R , Eqn. (19), and σ_A , Eqn. (20), both depend on the cell cycle time, $t - t_0$, and generation time, T_G . For short mRNA and short protein lifetimes we obtain from Eqns. (27), (28) the following ratio:

$$\frac{\sigma_A^2}{\sigma_R^2} = B_{eq} \frac{\text{var}(A)}{\langle A \rangle} \frac{\lambda_R^+}{\gamma_A + \lambda_R^-} \left(1 + \frac{\lambda_R^-}{\gamma_A + \lambda_X^-} \right). \quad (32)$$

If we subsequently assume, as it has been done by Bar-Even et al. and Raser and O'Shea, that protein lifetimes are longer, such that switching between promoter states occurs more frequently than a protein degradation event, $\lambda_R^- \gg \gamma_A \gg \lambda_X^-$, Eqn. (32) reduces to Eqn. (31). It follows from Eqns. (31) and (32) that for fixed rates the probability of activator binding, B_{eq} , determines the value of the ratio σ_A^2/σ_R^2 , i.e. the induction level of the activator selects the predominant source of noise. Thus, we expect for highly expressed genes, $B_{eq} \rightarrow 1$, to show signature of noise from promoter activation provided $\lambda_A^- \lambda_R^+ > \gamma_A^2$ since in this case we find $\sigma_A^2/\sigma_R^2 > 1$. In contrast, we expect for low induced genes, $B_{eq} \ll 1$, to show signature of Poissonian noise from mRNA synthesis, since in this case $\sigma_A^2/\sigma_R^2 < 1$ holds given $\lambda_A^- \lambda_R^+/\gamma_A^2$ is not too large. Therefore, the induction level of the activator, B_{eq} , provides an excellent explanation for the observation of different noise contributions even in the same organism. In fact, activator induction is expected to be quite low for experiments with native genes (9; 14). Bar-Even et al. investigated native genes implying a large set of low induced genes, $B_{eq} \ll 1$, such that transcription is the prevailing source of noise. In contrast, Raser and O'Shea constructed yeast strains that expressed CFP and GFP proteins from identical promoters. In constructs activator induction is very high, $B_{eq} \rightarrow 1$, such that we expect that promoter activation noise is the dominant noise contribution. In Fig. 2 we present the mean protein abundance vs. noise. In order to select arbitrary time points t within a given cell cycle, $0 \leq t - t_0 \leq T_G$, we use the full expressions for the mean amount of protein, Eqn. (15), and noise contributions (see SI), because the approximations presented in Eqns. (17), (19) and (20) are only valid in the asymptotic limit $t - t_0 \gg (\lambda_R^-)^{-1}$. We calculate the mean amount of protein and the noise contributions for several genes at randomly selected time points for several induction levels of the activator binding B_{eq} . In Fig. 2A we assume a low induction of the activator, where the mean induction level B_{eq}

equals 0.07, and therefore mimics the experimental set-up of Bar-Even et al. The noise contribution arising from transcription (blue circles) dominates the overall noise (magenta diamonds). In Fig. 2B, however, noise from promoter activation (green squares) overrules noise from transcription. This can be arranged with an highly induced activator, with mean induction level $B_{eq} = 0.7$, which mimics the experimental set-up of Raser and O’Shea. We conclude that both experimental scenarios can be qualitatively reproduced very well with our stochastic model by varying the induction level of the activator binding.

Activator binding and promoter activation determine population distribution

In eukaryotes, promoter activation is believed to occur due to chromatin remodelling (5) which erratically uncovers transcription-factor binding sites. Activator binding, however, is assumed to be quite fast and frequent, because of the high copy number of TFs. Therefore, independence of activator binding and promoter activation seems to be a reasonable assumption in eukaryotes. For prokaryotes, the situation is less clear since a possible explanation or mechanism for the slow process of promoter activation is still lacking, although it has been measured quite accurately (21; 22). In recent experiments Elf et al. measured the time scale for the binding/unbinding of an activator/TF at the single-molecule level in a living cell of *E. coli*. The binding/unbinding of highly abundant TFs is also suggested to be quite fast (33). Therefore, activator binding does not seem to be the limiting step within the process of gene expression. In order to gain insight into the influence of the activator binding on promoter activation in prokaryotes, we compare experimental data with simulations. Kollmann et al. (see (35), Fig. 2a; redrawn in Fig. 3A, inset), compared the mean expression of CheY in a wild-type of *E. coli* and *flgM* cells, where the upstream transcription inhibitor, FlgM, was deleted. The deletion corresponds to a sevenfold over expression of CheY. Several effects of an activator/repressor on the activation of the promoter are possible. We discuss the three most intuitive scenarios:

1. *Activator binding and promoter activation are independent:* The RNA-polymerase can start transcription if and only if the activator is bound (repressor is unbound) *and* the promoter is active.
2. *Activator binding enhances promoter switch-on rate:* For an experimentally observed sevenfold over expression we assume that the switch-on rate λ_A^+ of the over expression of

CheY is enhanced compared to that of the wild-type. The switch-off rate λ_A^- is not affected.

3. Activator binding decreases promoter switch-off rate: For an experimentally observed sevenfold over expression we assume that the switch-off rate λ_A^- of the over expression is decreased compared to that of the wild-type. The switch-on rate λ_A^+ is not affected.

Of course, combinations of the mentioned scenarios are possible and likely to occur in nature. However, to keep the estimated parameters identifiable, we only focus on these three limiting scenarios.

The experimental data shows that the mean protein level of the over expressing *flgM* cells is sevenfold higher than that of the wild-type cells. Furthermore, the standard deviation of the population distribution for the *flgM* cells increases quite significantly compared to the wild-type cells (Fig. 3A (inset) and Table 1). Fig. 3A shows the population distributions of the wild-type and *flgM* cells for the different scenarios after parameter estimation (see Material and Methods). The parameters are estimated such that the wild-type standard deviation and the mean fluorescence level of the *flgM* cells are represented best. The estimated parameters (B_{eq} for the first scenario, λ_A^+ , λ_A^- and the level of the over expression OE for second and third scenario) and the corresponding characteristics of the population distributions are summarized in Table 1.

The simulations reveal that there exists a set of parameters for the first scenario, where the activator binding does not influence the promoter activation process directly, such that the characteristic standard deviation of the wild-type and the mean fluorescence level of the *flgM* cells is reproduced very well ($\chi^2 = 0.004$). Furthermore, it also mimics (without any optimization) the increased standard deviation of the *flgM* cells (cf. Table 1 and Fig. 3A, red line denotes wild-type, black line denotes sevenfold over expression). The mean activator binding for the *flgM* cells is 7.3 times larger than that of the wild-type cells ($B_{eq,WT} = 0.13$, $B_{eq,OE} = 0.95$) which leads to an about sevenfold protein over expression of the mean fluorescence level. For the second scenario, where activator binding and promoter activation are not independent of each other, but activator binding enhances the promoter switch-on rate, the simulations with the estimated parameters does not represent the characteristic standard deviation of the wild-type and the mean fluorescence level of the *flgM* cells equivalently well ($\chi^2 = 0.67$, Fig. 3A green and blue lines denote wild-type and *flgM* cells, resp.). The standard deviation of the *flgM* cells becomes much larger than the experimental one. The estimated parameters are given by $\lambda_A^+ = 0.005$ for the wild-type

switch-on rate and $OE = 413$ for the increased multiplication factor for the *flgM* cells. The promoter is switched-on in 5% of the time for wild-type cells and 95% of the time for *flgM* cells. For the third scenario, where the binding of the activator decreases the promoter switch-off rate, a set of parameters can be found such that both characteristics are reproduced ($\chi^2 = 0.04$, Fig. 3A cyan and magenta dashed lines represent wild-type and *flgM* cells, resp.). The standard deviation of the *flgM* cells is also increased quite well. The resulting parameters $\lambda_A^- = 0.39$ for the wild-type switch-off rate and $OE = 924$ for the reduction factor for the *flgM* cells imply that the promoter is switched-off 90% of the time for wild-type and 1% of the time for *flgM* cells. If we compare the resulting skewness of the *flgM* cells in each scenario with the experimentally measured one we find that the second scenario has the most positive skewness, but this scenario does not fit the characteristics quite well. For the first scenario, the characteristics are represented very well and the skewness is also increased compared to the third scenario. We conclude that activator binding and promoter activation can be considered to good approximation as independent processes in prokaryotes. Of course, this hypothesis has to be investigated in further experiments.

Recently, stochastic dynamics has been linked to population distributions implying the classical model of burst-like transcription and translation (31). A Gamma-distribution fitted the stochastic simulations well and reproduced specific shapes of the population distribution at steady-state.

Here, we investigate the influence of the ratio $\kappa = \lambda_A^+/\lambda_A^-$ of the promoter switching rates on the shape of the population distribution. Simulations of the model presented in Fig. 1 reveal that different cell distribution shapes can be generated for a fixed switch-off λ_A^- rate by varying the switch on rate λ_A^+ . If the promoter activation is much smaller than its inactivation, $\kappa \ll 1$, the resulting protein distribution peaks at zero (Fig. 3B, blue line). For appropriate promoter activation rate, the protein distribution seems to be log-normally distributed (green line). If the promoter activation rate exceeds the inactivation rate, $\kappa \geq 1$, the protein distribution will be shifted to a normal distribution (magenta). The inset of Fig. 3B shows the corresponding protein trajectories of the different switch-on rates. The remaining reaction rates of activator binding, transcription and translation have an impact on the mean amount of protein or the noise strength, but not on the shape of the protein

distribution itself. If, however, the time scale of promoter activation is much faster than that of activator binding, e.g., for strongly repressed genes, the shape of the distribution can also be influenced by strong fluctuations in the activator concentration, i.e., the roles of the activator binding and promoter activation are interchanged. Thus, the results of Fig. 3B can also be obtained by fixing the promoter to its *on* state and choosing appropriate parameters for the activator binding.

Discussion

Different experiments have identified different causes for the main contribution to gene expression noise. This implies that there might be no general rule for the main source of noise or comprehensive knowledge of the overall noise architecture. However, the theoretical framework presented in this study allows to quantify the relative contributions of the different sources of gene expression noise. Taking cell division explicitly into account we derive time-dependent, non-equilibrium solutions for the mean amount of mRNA and protein as well as for the noise contributions from promoter activation and transcription. In order to interpret the analytical results with respect to their biological relevance, we discuss asymptotic cases for the molecule lifetimes representing essential physiological cases. Our analysis confirms the intuition that molecule lifetimes, compared to generation time, determine the influence of the noise contributions from previous generations on the overall noise level of the actual generation. Long-term memory effects and noise accumulation from previous generations might become important if molecule lifetimes are much larger than the generation time. In addition, the main contribution to the cell-to-cell variation within a clonal population depends strongly on the kinetic rates associated with the expression of each individual gene. We show that the induction level of the activator or TF binding, B_{eq} , determines crucially the ratio of noise from promoter activation to noise from transcription, σ_A^2/σ_R^2 , and thus the dominant source of noise. It follows that the experimental set-up for the same organism *S. cerevisiae* of *in-vivo* experiments performed by Bar-Even et al. (7) and *in-vitro* experiments by Raser and O'Shea (5), plays a fundamental role for the experimentally measured noise level. Low induced genes (7) bear clear transcriptional noise signature due to low-copy number of mRNA molecules, whereas highly induced genes (5) show typical characteristics of noise stemming from promoter activation (Fig. 2). Therefore,

the proposed model provides an explanation for these contradictory experimental results since it is able to reproduce both results for the same organism *S. cerevisiae*. Additionally, this suggests that the incorporation of an activator binding process acting independent of the promoter activation process is important for a theoretical description of gene expression. In eukaryotes, independence of activator binding and promoter activation is a reasonable assumption whereas in prokaryotes the situation is less clear. Based on published, experimentally measured cell distribution of wild-type and *flgM* cells of *E. coli*, we performed parameter estimation with the proposed model to discriminate between limiting cases of the effect of activator binding on promoter activation. Simulations reveal that for the first and third scenario a set of parameters can be found such that the characteristic standard deviation of the wild-type cells and mean fluorescence level of the *flgM* cells can be reproduced quite well. An increased standard deviation for the *flgM* cells can also be observed in both scenarios such that a qualitative distinction between both scenarios based on simulations seems to be difficult. The biological interpretation of the estimated parameters, however, suggest that the first scenario, i.e., the independence of promoter activation and activator binding, is more likely since the cellular effort of a sevenfold increase of the mean activator binding B_{eq} is reasonable. If the binding of the activator increases/reduces the promoter switch-on/-off rate, in the second and third scenario, respectively, the promoter is switched-off most of the time (95%/90%) for the wild-type. However, these delay times for the protein production, even for a repressed gene, contradict the experimental observations in *E. coli* where proteins are produced quite continuously (pers. communication V. Sourjik). Since the protein level should be strongly optimized to allow for reliable information processing, it seems to be very unlikely that promoters have evolved which are so strongly repressed that they are switched-off nearly all of the time. Furthermore, the over expression factor OE is estimated to be very large in the second ($OE = 413$) and third ($OE = 924$) scenario. This implies that an experimental depletion of an upstream inhibitor leading to a sevenfold increased mean fluorescence level corresponds to a theoretical $\approx 400/900$ fold increase/decrease of the switch-on/-off rate. The cellular effort to achieve this is expected to be very high, such that these scenarios seem to be inefficient.

The increased skewness (skewness of the *flgM* cells of about 1.74) observed by Kollmann et al. in experiments is not reflected in any scenario of our underlying model. The proposed

model is a reduced description of the overall system and does not take into account the complex flagella network such that an entire coincidence of the experiments and simulations is not expected. The skewness might also be influenced by external factors, like variations in ribosome or polymerase concentrations or by feedbacks of downstream to upstream processes, but none of these features are explicitly included in the present model.

Thus, the first scenario reproduces the experimental observations quite well and the estimated parameters can also be interpreted biologically reasonable. This observation implies that the promoter activation rates might be an intrinsic property of the biological system, both for the wild-type and the over expression line. One way to regulate gene expression is to finetune and control the mean binding of the activator.

If the time scales of activator binding and promoter activation can be separated, the ratio between the promoter activation and inactivation rate determines the shape of the population distribution (Fig. 3B). This suggests that the often interpreted log-normal distribution of single cells does not result from consecutively multiplicative stochastic processes but rather reflects the switching rates of the slowest process within gene expression.

Materials and Methods

Simulating stochastic processes. We assume that activator binding and promoter activation can be described by a random telegraph process with transition rates λ_B^+ , λ_B^- (activator binding), λ_A^+ and λ_A^- (promoter activation). The initial state of the promoter is determined by drawing a uniformly distributed random number (URN) $r \in [0, 1]$ and checking whether $r < \lambda_A^+/\gamma_A$, such that the promoter is *on*. Otherwise it is *off* in its initial state. Transcription is a *birth-and-death* process with time-dependent synthesis rate $\lambda_R^+(t) = \lambda_R^+ A(t) B(t)$, i.e. mRNA can only be synthesized if the activator is bound *and* the promoter is in its *on*-state (see Results and Fig. 1). The original Gillespie-algorithm (36) has been refined (for review see (37)), but also modified and extended to model growing cell volume via time-dependent reaction rates (38). To determine the next time τ of the reaction and the next reaction μ itself for time-dependent reaction rates, we follow along the lines of Gillespie (36) and Lu et al. (38) and arrive at the cumulative distribution function

$$F(\tau) = 1 - \exp \left[- \sum_{\mu} \int_0^{\tau} a_{\mu}(t + \tau') d\tau' \right] =: 1 - P_0(\tau). \quad (33)$$

Drawing a URN $\bar{u}_1 \in [0, 1]$ we set $\bar{u}_1 = 1 - P_0(\tau)$ and obtain, since $1 - \bar{u}_1$ is also a URN, the new URN $u_1 = P_0(\tau)$. The stochastic time τ for the next reaction to occur is obtained by inverting this equation. We formulate the cumulative distribution function for transcription with a time-dependent synthesis rate $a_1(t) = \lambda_R^+(t) = \lambda_R^+ A(t)B(t)$ of mRNA. The degradation does not depend on time $\tau > t$ with transition rate $a_2 = \lambda_R^- R(t)$ where $R(t)$ represents the actual amount of mRNA at time t . We find that at time t , the next stochastic time τ has to satisfy

$$\ln(u_1) = -\lambda_R^+ \int_0^\tau B(t + \tau') A(t + \tau') d\tau' - \lambda_R^- R(t) \tau. \quad (34)$$

Drawing a second URN $u_2 \in [0, 1]$, the next reaction μ must fulfill the inequality

$$\lambda_R^+ A(t + \tau) B(t + \tau) < u_2 (\lambda_R^+ A(t + \tau) B(t + \tau) + \lambda_R^- R(t + \tau)) \quad (35)$$

The modified Gillespie-algorithm with time-dependent reaction rate $a_1(t)$ determines the next stochastic time τ in Eqn. (34) as the upper bound of the integral. In our case, however, the integrand has a very special form, i.e., it is 1 if and only if the activator is bound as well as the promoter is *on*. Otherwise the integrand is 0. Therefore, the integration becomes a simple summation over all *on*-states, O_τ , of the product of the activator times the promoter within the time interval $[t, t + \tau]$. We define α_τ as the ratio of the *on*-states to the time interval $[t, t + \tau]$, i.e., $0 \leq \alpha_\tau = O_\tau / \tau \leq 1$. The next time τ can therefore be calculated according to Eqn. (34) which reduces for an exponentially distributed stochastic variable $\ln(u_1) =: -z$ to

$$\tau = \frac{z}{\lambda_R^+ \alpha_\tau + \lambda_R^- R(t)}. \quad (36)$$

Note that α_τ depends on the single realization of $A(t)B(t)$ and is thus also a stochastic variable. If $A(t)$ and $B(t)$ are time-independent, e.g., $A(t)B(t) \equiv 1$, it follows that $\alpha_\tau = 1$ and the original Gillespie-algorithm (36) is recovered. Therefore, we use the original Gillespie-algorithm to determine the next time τ and reaction μ and check afterwards whether the proposed Gillespie-step can be performed or not. If $A(t + \tau)B(t + \tau) = 1$, mRNA synthesis can be realized, but if $A(t + \tau)B(t + \tau) = 0$ and mRNA synthesis is selected as reaction μ , the step is rejected and new URNs are drawn. In general, this procedure will always select a stochastic time τ which is smaller than that of the modified algorithm of Eqn. (36) since $0 \leq \alpha_\tau \leq 1$. However, the above procedure of taking the original Gillespie-algorithm and rejecting specific reactions is equivalent to the determination of the next time τ via Eqn. (36). To obtain the same stochastic time for both procedures, the following equation should hold:

$$z = \frac{\lambda_R^+ \alpha_\tau + \lambda_R^- R(t)}{\lambda_R^+ + \lambda_R^- R(t)} \bar{z}; \quad (37)$$

where z and \bar{z} are exponentially distributed variable stemming from a URN u via $\ln(u)$. This equation is true since an exponentially distributed variable can be described by the product of a constant (given a specific α_τ) times another exponentially distributed stochastic variable \bar{z} . Therefore, taking the original Gillespie-algorithm and rejecting specific reactions according to the time-dependent trajectory $A(t)B(t)$ is just another realization of the modified Gillespie from Eqn. (36) and averaging over a lot of trajectories yields the same result.

The generated mRNA trajectory $R(t)$ can directly be used to calculate the appropriate mean protein number $\langle X(t) \rangle_X$ given by Eqn. (8). At the end of one cell cycle, the cell divides symmetrically into two daughter cells. The mother's amount of protein and mRNA is divided binomially to both daughter cells. To demonstrate that the modified Gillespie-algorithm computes the correct solution, Fig. S1 shows a comparison of simulated mRNA and protein trajectories with the analytical ones from Eqs. (15) and (16), respectively.

Generating population distributions. In order to avoid dependence on the initial conditions, we start with a certain amount of mRNA and proteins and simulate in total 15 generations. After five generations we randomly determine one cell with its amount of mRNA and protein to be the mother cell for the next five generations. The system is equilibrated and the actual simulations can be started. We randomly choose one cell to be the mother cell and generate 10 offspring generations (=1024 cells in the 10th generation). To compare the influence of the length of the generation time T_G on the population distribution, we simulated populations with fixed generation time and varying generation time (choosing $T_G \in \mathcal{N}(\mu, \sigma)$). We did not observe any significant differences between these two scenarios and therefore fix for simplicity the generation time in all simulations.

Effect of activator binding on promoter activation: parameter estimation. A least-square fit is performed with MATLAB such that, after data normalization to mean wild-type fluorescence of 1, the experimental standard deviation of the wild-type cells ($\sigma = 0.69$) and mean *flgM* fluorescence ($\mu = 6.96$) of the population distributions are best represented. For each optimization step, $20 \times 2^9 = 10240$ realizations of the proposed model are generated. The following parameters are fixed for the simulations: $\lambda_R^+ = 2$, $\lambda_R^- = 0.2$, $\lambda_X^+ = 4$ and $\lambda_X^- = 10^{-4}$. In the first scenario, the promoter switch-on/-off rates are set to $\lambda_A^+ = 0.05$ and $\lambda_A^- = 0.1$, representing realistic kinetic rates for a repressed gene (21). The mean activator binding rates of wild-type ($B_{eq,WT}$) and *flgM* cells ($B_{eq,OE}$) are estimated separately. For the second and third scenario we assume a mean activator

binding of $B_{eq} = 0.5$ for the wild-type as well as for the *flgM* cells. The promotor switch-off rate is set to $\lambda_A^- = 0.1$ and the promotor switch-on $\lambda_{A,WT}^+$ for the wild-type and the strength of the over expression, OE , are estimated in the second scenario. If the activator is bound, the promotor switch-on rate for the *flgM* cells, $\lambda_{A,OE}^+$, will be enhanced by the over expression factor OE , i.e. $\lambda_{A,OE}^+ = OE \times \lambda_{A,WT}^+$. In the third scenario we set the promotor switch-on rate to $\lambda_A^+ = 0.05$ and estimate the promotor switch-off rate $\lambda_{A,WT}^-$ for the wild-type and the strength of the over expression OE . If the activator is bound, the promotor switch-off rate for the *flgM* cells, $\lambda_{A,OE}^-$, will be reduced by the over expression factor OE , i.e. $\lambda_{A,OE}^- = \lambda_{A,WT}^-/OE$. The estimated parameters are summarized in Table 1.

Acknowledgements

This work was financially supported by the Graduate School “Plant Signal Systems” (DFG grant no. GK1305 for JR) and the DFG Emmy Noether-Program (for MK).

Competing interests The authors have declared that no competing interests exist.

Author’s contributions MK initiated and supervised the study, MK and JR carried out the analytical calculations, JR performed the simulations, JR, CF and MK wrote the manuscript with useful comments from JT.

Abbreviations G/C/YFP, green/cyan/yellow fluorescent protein; TF, transcription factor; OE, over expression; URN, uniform random number

References

- [1] Spudich JL and Koshland, DE Jr. (1976) Non-genetic individuality: chance in the single cell. Nature 262: 467-71.
- [2] Raser JM and O’Shea EK (2005) Noise in gene expression: Origins, consequences, and control. Science 309: 2010-3.
- [3] Avery SV (2006) Microbial cell individuality and the underlying sources of heterogeneity. Nat Rev Microbiol 4: 577-87.
- [4] Elowitz MB, Levine AJ, Siggia ED and Swain PS (2002) Stochastic gene expression in a single cell. Science 297: 1183-6.

- [5] Raser JM and O'Shea EK (2004) Control of stochasticity in eukaryotic gene expression. *Science* 304: 1811-4.
- [6] Volfson D, Marciniak J, Blake WJ, Ostroff N, Tsimring LS and Hasty J (2006) Origins of extrinsic variability in eukaryotic gene expression. *Nature* 439: 861-4.
- [7] Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea EK, Pilpel Y and Barkai N (2006) Noise in protein expression scales with natural protein abundance. *Nat Gen* 38: 636-43.
- [8] Becskei A, Kaufmann BB and van Oudenaarden A (2005) Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat Gen* 37: 937-44.
- [9] Raj A, Peskin CS, Tranchina D, Vargas DY and Tyagi S (2006) Stochastic mRNA synthesis in mammalian cells. *PLoS Biol* 4: 1707-19.
- [10] Blake WJ, Koern M, Cantor CR and Collins JJ (2003) Noise in eukaryotic gene expression. *Nature* 422: 633-7.
- [11] Pedraza JM and van Oudenaarden A (2005) Noise propagation in gene networks. *Science* 307: 1183-6.
- [12] Golding I, and Cox EC (2006) Eukaryotic transcription: what does it mean for a gene to be 'on'? *Curr Biol* 16: R371-3.
- [13] Mettetal JT, Muzzey D, Pedraza JM, Ozbudak EM and van Oudenaarden A (2006) Predicting stochastic gene expression dynamics in single cells. *Proc Natl Acad Sci U S A* 103: 7304-9.
- [14] Newman JR, Ghaemmighami S, Ihmels J, Breslow DK, Noble M, DeRisi JL and Weissman JS (2006) Single-cell proteomic analysis of *S.cerevisiae* reveals the architecture of biological noise. *Nature* 441: 840-6.
- [15] Sigal A, Milo R, Cohen A, Geva-Zatorsky N, Klein Y, Liron Y, Rosenfeld N, Danon T, Perzov N and Alon U (2006) Variability and memory of protein levels in human cells. *Nature* 444: 643-6.
- [16] Rosenfeld N, Young JW, Alon U, Swain PS and Elowitz MB (2005) Gene regulation at the single-cell level. *Science* 307: 1962-5.
- [17] van Kampen, N G (1992) *Stochastic Processes in Physics and Chemistry*. North Holland. Rev. and enlarged edition
- [18] Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, Endy D, Pesce CG and Brent R (2005) Regulated cell-to-cell variation in a cell-fate decision system. *Nature* 437:699-706
- [19] Kaufmann BB and van Oudenaarden A (2007) Stochastic gene expression: from single

- molecules to the proteome. *Curr Opin Genet Dev* 17: 1-6
- [20] Ozbudak EM, Thattai M, Kurtser I, Grossman AD and van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. *Nat Gen* 31: 69-73.
 - [21] Golding I, Paulsson J, Zawilski SM and Cox EC (2005) Real-time kinetics of gene activity in individual bacteria. *Cell* 123: 1025-36.
 - [22] Cai L, Friedman N and Xie XS (2006) Stochastic protein expression in individual cells at the single molecule level. *Nature* 440: 358-62.
 - [23] Balaban NQ, Merrin J, Chait R, Kowalik L and Leibler S (2004) Bacterial persistence as a phenotypic switch. *Science* 305: 1622-5.
 - [24] Suel GM, Garcia-Ojalvo J, Liberman LM and Elowitz MB (2006) An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* 440: 545-50.
 - [25] Berg OG (1978) A model for statistical fluctuations of protein numbers in a microbial population. *J Theo Biol* 71: 587-603
 - [26] Thattai M and van Oudenaarden A (2001) Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci U S A* 98: 8614-9.
 - [27] Swain PS, Elowitz MB and Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci U S A* 99: 12795-800.
 - [28] Swain PS (2004) Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *J Mol Biol* 344: 965-76.
 - [29] Paulsson J (2004) Summing up the noise in gene networks. *Nature* 427: 415-8.
 - [30] Paulsson J (2005) Models of stochastic gene expression. *Phys Life Rev* 2: 157-75.
 - [31] Friedman N, Cai L and Xie XS (2006) Linking Stochastic dynamics to population distribution: An analytical framework to gene expression. *Phys Rev Lett* 97: 168302.
 - [32] Alon U (2006) *An Introduction to Systems Biology: Design principles of biological circuits.* Chapman & Hall/CRC. 301 p.
 - [33] Elf J, Li GW and Xie XS (2007) Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* 316: 1191-4
 - [34] Tănase-Nicola S, Warren PB and ten Wolde PR (2006) Signal detection, modularity and the correlation between extrinsic and intrinsic noise in biochemical networks. *Phys Rev Lett* 97: 068102-4.
 - [35] Kollmann M, Løvdo L, Bartholomé K, Timmer J and Sourjik V (2005) Design principles of

- a bacterial signalling network. *Nature* 438: 504-7
- [36] Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81(25): 2340-2361
- [37] Gillespie DT and Petzold LR (2006) Numerical simulation for biochemical kinetics. In: Szallasi, Z, Stelling, J, Perival, V editors. *Systems Modeling in Cellular Biology*, MIT press, pp. 331-353
- [38] Lu T, Volfson D, Tsimring L and Hasty J (2004) Cellular growth and division in the Gillespie algorithm. *Syst Biol* 1(1):121-128

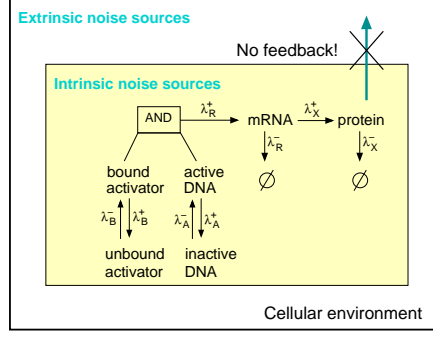


FIG. 1 Definition and reaction scheme of single-gene reporter system (yellow box) within intracellular environment (large box). Intrinsic and extrinsic noise can only be distinguished if expression level of reporter system does not influence extrinsic factors. Transition rates are defined in the text.

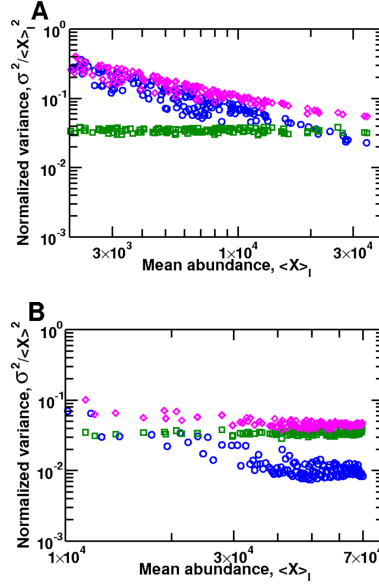


FIG. 2 Simulated mean abundance vs. noise for different genes. Transcriptional contribution (blue circles), noise from promoter activation (green squares) and overall noise (magenta diamonds).

A) Low induced activator, $B_{eq} = 0.07$, leads to $\sigma_A^2/\sigma_R^2 < 1$ such that transcriptional noise dominates. **B)** Highly induced activator, $B_{eq} = 0.7$, leads to $\sigma_A^2/\sigma_R^2 > 1$ such that noise from promoter activation dominates.

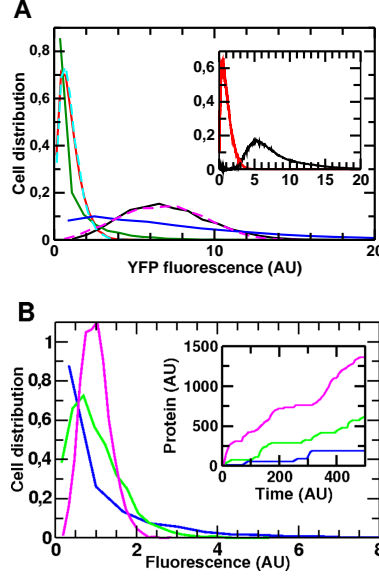


FIG. 3 **A)** Simulations of different effects of activator binding on promoter switch-on/off rates for fixed and estimated parameters. Red and black: wild-type and *flgM* cells of assuming that promoter activation and activator binding are independent (first scenario); green and blue: wild-type and *flgM* cells assuming that the activator binding enhances promoter switch-on rate (second scenario); cyan and magenta dashed lines: wild-type and *flgM* cells assuming activator binding decreases promoter switch-off rate (third scenario). Means, standard deviations and estimated parameters are summarized in Table 1, fixed parameters are found in the Materials and Methods. The inset shows the experimental levels of CheY, expressed as YFP fusion from native chromosomal position for wild-type (red) and *flgM* cells (black). Redrawn from (35). **B)** Different population distributions (*case I*), each normalized to mean 1 for varying promoter switch-on rates λ_A^+ at fixed switch-off rate $\lambda_A^- = 0.1$. Mean activator binding: $B_{eq} = \frac{1}{7}$. Blue: $\lambda_A^+ = 0.01$; green: $\lambda_A^+ = 0.05$; magenta: $\lambda_A^+ = 1$. The inset shows corresponding simulated protein trajectories for each switch-on rate λ_A^+ .

		μ	σ	skewness	estimated parameter	χ^2
Experiment	wild-type	1	0.69	2.51		
	<i>flgM</i> cells	6.96	3.38	1.74		
Simulation	wild-type	1	0.67	1.26	$B_{eq} = 0.13$	0.004
1st scenario	<i>flgM</i> cells	7.02	2.63	0.44	$B_{eq} = 0.94$	
Simulation	wild-type	1	1.51	2.86	$\lambda_A^+ = 0.005$	0.67
2nd scenario	<i>flgM</i> cells	6.92	5.25	1.07	$OE = 413$	
Simulation	wild-type	1	0.63	1.17	$\lambda_A^- = 0.39$	0.04
3rd scenario	<i>flgM</i> cells	6.77	2.56	0.1	$OE = 924$	

TABLE 1 Characteristic mean (μ), standard deviation (σ) and skewness for population distributions from experiments of Kollmann et al. (35) and simulations using the proposed stochastic model. Cells in which the upstream transcription inhibitor, the anti-sigma factor FlgM, was deleted, are denoted by *flgM* cells. The deletion leads to a sevenfold over expression. Differences in the simulated scenarios, interpretation of the factor *OE* and parameter estimation are described in the text and in Material and Methods.

Supporting Information

I. DERIVATION OF TIME-DEPENDENT SOLUTION FOR TRANSCRIPTION

For an activated promoter, the process of transcription is assumed to be Poissonian. Assume that $R(t)$ is the number of mRNA transcribed from an activated promoter $A(t)B(t) = 1$. The stochastic variables $A(t), B(t) \in \{0, 1\}$ denote the promoter activation and activator binding, respectively. We assume $R(t) \xrightarrow{\lambda_R^+} R(t) + 1$ with time-dependent transition rate $\lambda_R^+(t) = \lambda_R^+ A(t)B(t)$ and for the inverse process $R(t) \xrightarrow{\lambda_R^-} R(t) - 1$ where λ_R^- is the degradation rate.

Assuming initially R_0 molecules at time t_0 , the stochastic Master equation for this Poisson process has the form

$$\begin{aligned} \partial_t P(R, t | R_0, t_0) = & \lambda_R^+(t) P(R-1, t | R_0, t_0) + (R+1) \lambda_R^- P(R+1, t | R_0, t_0) \\ & - (R \lambda_R^- + \lambda_R^+(t)) P(R, t | R_0, t_0). \end{aligned} \quad (\text{S1})$$

To solve the equation we define the generating function

$$\begin{aligned} G(s, t | R_0, t_0) &= \sum_R s^R P(R, t | R_0, t_0) \text{ such that} \\ \partial_t G(s, t | R_0, t_0) &= \lambda_R^+(t) (s-1) G(s, t | R_0, t_0) - \lambda_R^- (s-1) \partial_s G(s, t | R_0, t_0). \end{aligned} \quad (\text{S2})$$

We choose the ansatz

$$G(s, t) = e^{(s-1)\alpha(t)} \psi(s, t) \quad (\text{S3})$$

for which we obtain using (S2)

$$\begin{aligned} \partial_t G(s, t) &= \partial_t \alpha(t) (s-1) e^{(s-1)\alpha(t)} \psi(s, t) + e^{(s-1)\alpha(t)} \partial_t \psi(s, t) \\ &= \lambda_R^+(t) (s-1) e^{(s-1)\alpha(t)} \psi(s, t) - \lambda_R^- (s-1) e^{(s-1)\alpha(t)} [\alpha(t) \psi(s, t) + \partial_s \psi(s, t)] \\ &= (\lambda_R^+(t) - \lambda_R^- \alpha(t)) (s-1) e^{(s-1)\alpha(t)} \psi(s, t) - \lambda_R^- (s-1) \partial_s \psi(s, t) e^{(s-1)\alpha(t)}. \end{aligned} \quad (\text{S4})$$

Comparison of the coefficients yields

$$\partial_t \alpha(t) = \lambda_R^+(t) - \lambda_R^- \alpha(t) \quad (\text{S5})$$

$$\partial_t \psi(s, t) = -\lambda_R^- (s-1) \partial_s \psi(s, t). \quad (\text{S6})$$

The solution of (S5) is given by

$$\alpha(t) = \alpha(t_0)e^{-\lambda_R^-(t-t_0)} + \int_{t_0}^t e^{-\lambda_R^-(t-t')} \lambda_R^+(t') dt'. \quad (\text{S7})$$

To determine $\psi(s, t)$, we introduce a new variable $z := \ln(s - 1)$ (with $dz = \frac{1}{s-1}ds$), substitute s by z in (S6) and obtain the differential equation

$$\partial_t \tilde{\psi}(z, t) = -\lambda_R^- \partial_z \tilde{\psi}(z, t). \quad (\text{S8})$$

Its solution is an arbitrary function F of the variable $\lambda_R^- t - z$, so we can write $\tilde{\psi}(z, t) = F[e^{z-\lambda_R^- t}]$ and therefore (re-substituting z by s)

$$\psi(s, t) = F[(s - 1)e^{-\lambda_R^- t}]. \quad (\text{S9})$$

It follows for the generating function

$$G(s, t) = F[(s - 1)e^{-\lambda_R^- t}]e^{(s-1)\alpha(t)}. \quad (\text{S10})$$

Normalization requires $G(1, t|R_0, t_0) = 1$ and therefore $F[0] = 1$.

For a complete determination of F , we have to include the initial condition $P(R, 0|R_0, 0) = \delta_{RR_0}$ which implies

$$G(s, 0|R_0, t_0) = \sum_R s^R \underbrace{P(R, 0|R_0, 0)}_{=\delta_{RR_0}} = s^{R_0} = F[s - 1]e^{(s-1)\alpha(0)}. \quad (\text{S11})$$

We arrive at $F[s - 1] = [(s - 1) + 1]^{R_0} e^{-(s-1)\alpha(0)}$ which gives for the substitution $x := s - 1$

$$F[x] = [x + 1]^{R_0} e^{-x\alpha(0)}. \quad (\text{S12})$$

Replacing F in (S10) with our fully determined F of (S12) we obtain

$$G(s, t|N, 0) = [(s - 1)e^{-\lambda_R^- t} + 1]^N \exp[(s - 1)(\alpha(t) - \alpha(0)e^{-\lambda_R^- t})] \quad (\text{S13})$$

To start at an arbitrary time t_0 and molecule number R_0 we now include the effect of cell division into the generating function Eqn. (S13) of the transcription process. It is obvious that the noise level strongly depends on the mean number of mRNAs. If there were many mRNA copies the noise from cell division would become smaller, such that the process of cell division would not account for the variability among daughter cells. Conversely, only a few copies of mRNA result in a large variability among daughter cells. Therefore this effect

has to be taken into account when focussing on gene expression noise.

Activator binding and promoter activation are expected to be fairly independent on cell division and the amount of protein is simply halved if the cell divides symmetrically as for most systems the copy number of proteins is $> 10^2$. Thus transcription is the only process where we have to account for cell division in the probability density function.

Assuming exactly one symmetric cell division with binomial distribution of the mRNA at time t_1 within $t_0 < t_2$, the expression for the conditional probability density function of mRNA synthesis and decay is given by

$$P(R_2^-, t_2 | R_0^+, t_0) = \sum_{R_1^- = 0}^{\infty} \sum_{R_1^+ = 0}^{R_1^-} P(R_2^-, t_2 | R_1^+, t_1) \binom{R_1^-}{R_1^+} 2^{-R_1^-} P(R_1^-, t_1 | R_0^+, t_0), \quad (\text{S14})$$

with R_1^+ and R_1^- is the amount of mRNA of the daughter and mother cell, respectively.

It can be easily seen via induction over the number of cell divisions, that the probability density function of mRNA synthesis and decay after n cell divisions at t_1, \dots, t_n reads with $t_{n+1} > t > t_n$ and $t_1 > t' > t_0$

$$\begin{aligned} P(R, t | R', t') &= \sum_{R_n^- = 0}^{\infty} \sum_{R_n^+ = 0}^{R_n^-} P(R, t | R_n^+, t_n) \binom{R_n^-}{R_n^+} 2^{-R_n^-} P(R_n^-, t_n | R', t') \\ &= \sum_{R_n^- = 0}^{\infty} \sum_{R_n^+ = 0}^{R_n^-} \cdots \sum_{R_1^- = 0}^{\infty} \sum_{R_1^+ = 0}^{R_1^-} P(R, t | R_n^+, t_n) \binom{R_1^-}{R_1^+} 2^{-R_1^-} P(R_1^-, t_1 | R', t') \\ &\quad \times \prod_{i=2}^n \binom{R_i^-}{R_i^+} 2^{-R_i^-} P(R_i^-, t_i | R_{i-1}^+, t_{i-1}). \end{aligned} \quad (\text{S15})$$

The generating function corresponding to one cell division at t_1 is given by

$$\begin{aligned} G(s_2, t_2 | R_0, t_0) &= \sum_{R_2, R_1^+, R_1^-} s_2^{R_2} P(R_2, t_2 | R_1^+, t_1) \underbrace{\binom{R_1^-}{R_1^+} 2^{-R_1^-} P(R_1^-, t_1 | R_0, t_0)}_{=P(R_2, t_2 | R_0, t_0)} \quad (\text{S16}) \\ &\stackrel{(\text{S13})}{=} \sum_{R_1^+, R_1^-} [(s_2 - 1)e^{-\lambda_R^-(t_2 - t_1)} + 1]^{R_1^+} \exp \left[(s_2 - 1) \int_{t_1}^{t_2} e^{-\lambda_R^-(t_2 - t')} \lambda_R^+(t') dt' \right] \\ &\quad \times \binom{R_1^-}{R_1^+} 2^{-R_1^-} P(R_1^-, t_1 | R_0, t_0) \end{aligned}$$

$$\begin{aligned}
&= \sum_{R_1^-} \sum_{R_1^+=0}^{R_1^-} [(s_2 - 1)e^{-\lambda_R^-(t_2-t_1)} + 1]^{R_1^+} \binom{R_1^-}{R_1^+} 2^{-R_1^-} \\
&\quad \times \exp \left[(s_2 - 1) \int_{t_1}^{t_2} e^{-\lambda_R^-(t_2-t')} \lambda_R^+(t') dt' \right] P(R_1^-, t_1 | R_0, t_0) \\
&= \sum_{R_1^-} \left[\underbrace{\frac{1}{2}(s_2 - 1)e^{-\lambda_R^-(t_2-t_1)} + 1}_{s_1} \right]^{R_1^-} \\
&\quad \times \exp \left[(s_2 - 1) \int_{t_1}^{t_2} e^{-\lambda_R^-(t_2-t')} \lambda_R^+(t') dt' \right] P(R_1^-, t_1 | R_0, t_0) \tag{S17}
\end{aligned}$$

$$\begin{aligned}
&\stackrel{(S13)}{=} \left[(s_1 - 1)e^{-\lambda_R^-(t_1-t_0)} + 1 \right]^{R_0} \exp \left[(s_2 - 1) \int_{t_1}^{t_2} e^{-\lambda_R^-(t_2-t')} \lambda_R^+(t') dt' \right] \\
&\quad \times \exp \left[(s_1 - 1) \int_{t_0}^{t_1} e^{-\lambda_R^-(t_1-t')} \lambda_R^+(t') dt' \right] \tag{S18}
\end{aligned}$$

$$\begin{aligned}
&= \left[\frac{1}{2}(s_2 - 1)e^{-\lambda_R^-(t_2-t_0)} + 1 \right]^{R_0} \exp \left[(s_2 - 1) \int_{t_1}^{t_2} e^{-\lambda_R^-(t_2-t')} \lambda_R^+(t') dt' \right] \\
&\quad \times \exp \left[\frac{1}{2}(s_2 - 1) \int_{t_0}^{t_1} e^{-\lambda_R^-(t_2-t')} \lambda_R^+(t') dt' \right]. \tag{S19}
\end{aligned}$$

By repeated application of the method above, i.e. assuming a finite number of cell divisions, the generating function is given by

$$\begin{aligned}
G(s, t | R_0, t_0) &= \left[2^{-D(t, t_0)} (s - 1) e^{-\lambda_R^-(t-t_0)} + 1 \right]^{R_0} \\
&\quad \times \exp \left[(s - 1) \int_{t_0}^t \lambda_R^+(t') e^{-\lambda_R^-(t-t')} 2^{-D(t, t')} dt' \right] \\
&=: (\tilde{x} + 1)^{R_0} K(s, t) \tag{S20}
\end{aligned}$$

where $D(t, t') = D(t', t)$ represents the number of cell divisions within the time interval $[t, t']$. The above formula is easily proven via induction over the number $D(t, t')$ of cell divisions within $[t, t']$.

We first compute the mean mRNA copy number, $\langle R(t) \rangle_R$, with the help of the generating function Eqn. (S20). Setting $t_0 = -\infty$, i.e. including an infinite number of cell divisions, we find that

$$G(s, t | R_0, -\infty) = \exp [(s - 1)\alpha(t)] \tag{S21}$$

with $\alpha(t)$ defined by

$$\begin{aligned}
\alpha(t) &= \int_{-\infty}^t \lambda_R^+(t') e^{-\lambda_R^-(t-t')} 2^{-D(t,t')} dt' \\
&= \int_{-\infty}^{t_0} \lambda_R^+(t') e^{-\lambda_R^-(t-t')} 2^{-D(t,t')} dt' + \int_{t_0}^t \lambda_R^+(t') e^{-\lambda_R^-(t-t')} 2^{-D(t,t')} dt' \\
&= e^{-\lambda_R^-(t-t_0)} 2^{-D(t,t_0)} \alpha(t_0) + \int_{t_0}^t \lambda_R^+(t') e^{-\lambda_R^-(t-t')} 2^{-D(t,t')} dt'. \tag{S22}
\end{aligned}$$

Note that $D(t, t') = D(t, t'') + D(t'', t')$ and that $\alpha(t)$ introduced here is a generalization of $\alpha(t)$ without cell division (cf.(S7)).

As Eqn. (S20) is the generating function of a Poisson process we obtain the mean of $R(t)$ by calculating $\partial_s G|_{s=1}$

$$\alpha(t) = \langle R(t) \rangle_R = \langle R(t_0) \rangle_R 2^{-D(t_0,t)} e^{-\lambda_R^-(t-t_0)} + \int_{t_0}^t \lambda_R^+(t') e^{-\lambda_R^-(t-t')} 2^{-D(t,t')} dt'. \tag{S23}$$

To calculate the variance we use a generating function $\Theta(s, t)$ which contains the joint probability function $P(R, t; R_0, t_0)$ representing the probability that at time t there are R molecules and at time t_0 there are R_0 molecules. Afterwards we determine the variances by

$$\langle R(t)R(t_0) \rangle_R = \partial_s \Theta|_{s=1}. \tag{S24}$$

We again assume that initially the process is Poissonian-distributed and use the notation of (S20), i.e. $G(s, t|R_0, t_0) = (x+1)^{R_0} K(s, t)$. The generating function $\Theta(s, t)$ is given by

$$\begin{aligned}
\Theta(s, t) &:= \sum_{R_0} R_0 \sum_R s^R P(R, t; R_0, t_0) = \sum_{R_0} R_0 \underbrace{\sum_R s^R P(R, t|R_0, t_0)}_{G(s,t|R_0,t_0)} P(R_0, t_0) \\
&= \sum_{R_0} R_0 (x+1)^{R_0} K(s, t) \frac{e^{-\alpha(t_0)} \alpha(t_0)^{R_0}}{R_0!} \\
&= \sum_{R_0} (x+1) K(s, t) \partial_x (x+1)^{R_0} \frac{e^{-\alpha(t_0)} \alpha(t_0)^{R_0}}{R_0!} \\
&= (x+1) K(s, t) \partial_x e^{-\alpha(t_0)} \underbrace{\sum_{R_0} \frac{(\alpha(t_0)(x+1))^{R_0}}{R_0!}}_{=e^{\alpha(t_0)(x+1)}} \\
&= (x+1) K(s, t) \alpha(t_0) e^{\alpha(t_0)x}
\end{aligned}$$

$$\begin{aligned}
&= [2^{-D(t_0,t)}(s-1)e^{-\lambda_R^-(t-t_0)} + 1]\alpha(t_0) \\
&\quad \times \exp \left[(s-1)(\alpha(t_0)2^{-D(t_0,t)}e^{-\lambda_R^-(t-t_0)} + \int_{t_0}^t \lambda_R^+(t')e^{-\lambda_R^-(t-t')}2^{-D(t,t')}dt') \right] \\
&= [2^{-D(t_0,t)}(s-1)e^{-\lambda_R^-(t-t_0)} + 1]\alpha(t_0) \exp[(s-1)\langle R(t) \rangle_R].
\end{aligned} \tag{S25}$$

It follows that

$$\langle R(t)R(t_0) \rangle_R = \partial_s \Theta|_{s=1} = \alpha(t_0)(2^{-D(t_0,t)}e^{-\lambda_R^-(t-t_0)} + \langle R(t) \rangle_R). \tag{S26}$$

Replacing t_0 by t' we obtain for the time correlation

$$\begin{aligned}
\langle R(t)R(t') \rangle_R &= \alpha(t')(2^{-D(t',t)}e^{-\lambda_R^-(t-t')} + \langle R(t) \rangle_R) \\
&= \langle R(t') \rangle_R 2^{-D(t',t)}e^{-\lambda_R^-(t-t')} + \langle R(t) \rangle_R \langle R(t') \rangle_R.
\end{aligned} \tag{S27}$$

II. PARTIAL CONTRIBUTIONS TO GENE EXPRESSION NOISE

For a one-gene system (cf. Fig. 1 main text) whose gene product does not interfere with external parameters (e.g. expression for genes forming ribosomes do infer with the translation rate and thus with external parameters) the total variation σ_{tot}^2 can be written as a sum of the following contributions

$$\sigma_{tot}^2 = \sigma_I^2 + \sigma_E^2. \tag{S28}$$

The term σ_I^2 is the intrinsic noise contribution for fixed environmental conditions whereas the term σ_E^2 represents the effect of extrinsic noise on the mean expression level. The above expression for intrinsic noise is defined by

$$\sigma_I^2 = \langle (X(t) - \langle X(t) \rangle_{X,R,A,B})^2 \rangle_{X,R,A,B} \tag{S29}$$

$$= \langle (X(t) + \langle X(t) \rangle_X - \langle X(t) \rangle_X - \langle X(t) \rangle_{X,R,A,B})^2 \rangle_{X,R,A,B} \tag{S30}$$

$$\begin{aligned}
&\stackrel{\langle \cdot \rangle_X}{=} \langle (X(t) - \langle X(t) \rangle_X)^2 \rangle_{X,R,A,B} + \langle (\langle X(t) \rangle_X - \langle X(t) \rangle_{X,R,A,B})^2 \rangle_{R,A,B} \\
&+ \underbrace{\langle 2(X(t) - \langle X(t) \rangle_X)(\langle X(t) \rangle_X - \langle X(t) \rangle_{X,R,A,B}) \rangle_{X,R,A,B}}_{=0, \text{ after } \langle \cdot \rangle_X}
\end{aligned} \tag{S31}$$

$$\begin{aligned}
&= \langle (X(t) - \langle X(t) \rangle_X)^2 \rangle_{X,R,A,B} + \langle (\langle X(t) \rangle_X - \langle X(t) \rangle_{X,R})^2 \rangle_{R,A,B} \\
&+ \langle (\langle X(t) \rangle_{X,R} - \langle X(t) \rangle_{X,R,A})^2 \rangle_{A,B} + \langle (\langle X(t) \rangle_{X,R,A} - \langle X(t) \rangle_{X,R,A,B})^2 \rangle_B
\end{aligned} \tag{S32}$$

$$= \sigma_X^2 + \sigma_R^2 + \sigma_A^2 + \sigma_B^2 \tag{S33}$$

where the methods of Eqns. (S30) and (S31) are applied iteratively to the remaining terms. The joint probability $P(X(t); R(t); A(t); B(t))$ can be written in terms of the following conditional probabilities, $P(X(t); R(t); A(t); B(t)) = P(X(t)|R(t))P(R(t)|A(t), B(t))P(A(t))P(B(t))$. The averaging over the processes is given by $\langle X(t) \rangle_X = \int X(t)P(X(t)|R(t))dX(t)$ where $\int P(X(t)|R(t))dX(t) = 1$.

The resulting terms of Eqn. (S33) give the second moment contributions from translation, σ_X^2 , transcription, σ_R^2 , promoter activation, σ_A^2 , and activator binding, σ_B^2 . Contribution from extrinsic noise consist of extrinsic effects on the mean expression and an increase of intrinsic noise due to extrinsic factors. The above contributions can be calculated analytically as shown in the following.

As mentioned in the main text, we assume a separation of the time scales of activator binding and promoter activation, such that we assume an equilibrated activator binding process. To make the calculations feasible, we additionally assume a fixed induction level of the activator, B_{eq} . We use the following abbreviations, where T_G is the generation time, α denotes either X or R , β represents either A or E and $n \in \mathbb{N}$.

$$Q_X = \frac{\lambda_X^+ \lambda_R^+}{\lambda_X^- \lambda_R^-} \langle A \rangle \quad (\text{S34})$$

$$\mathcal{A} = \langle A \rangle B_{eq} \lambda_R^+ \lambda_X^+ \quad (\text{S35})$$

$$Z_\alpha^{(n)}(T_G) = \frac{1 - \exp[-n\lambda_\alpha^- T_G]}{1 - 2^{-n} \exp[-n\lambda_\alpha^- T_G]} \quad (\text{S36})$$

$$Z_\beta^{(n)}(T_G) = \frac{1 - \exp[-n\gamma_\beta^- T_G]}{1 - 2^{-n} \exp[-n\gamma_\beta^- T_G]} \quad (\text{S37})$$

$$Z_{\alpha\beta}(T_G) = \frac{1 - \exp[(\lambda_\alpha^- + \gamma_\beta^+) T_G]}{1 - 2^{-1} \exp[(\lambda_\alpha^- + \gamma_\beta^+) T_G]}. \quad (\text{S38})$$

The expression levels within a large population are calculated after infinitely many cell divisions. The cell division times are given by t_i ; by $i = 0$ we define the last cell division that has occurred. Furthermore, we denote by t_i^- and t_i^+ the time just before and after the cell division, respectively. Given the n -th moment of the stochastic function for mRNA, $\mu_R^{(n)}(t)$, and protein number, $\mu_X^{(n)}(t)$, the following invariance conditions can be derived

$$\mu_R^{(n)}(t) = \mu_R^{(n)}(t + kT_G) \quad \forall \quad k \in \mathbb{Z} \quad (\text{S39})$$

$$\mu_X^{(n)}(t) = \mu_X^{(n)}(t + kT_G) \quad \forall \quad k \in \mathbb{Z}. \quad (\text{S40})$$

Before the next cell division at time t_1 , the mean mRNA and mean protein number of a large population is given by

$$\begin{aligned}\langle X(t) \rangle_{X,R,A} &= \langle \hat{X}(t - t_0^+) \rangle_{X,R,A} + \frac{1}{2} e^{-\lambda_X^-(t-t_0)} \langle X(t_0^-) \rangle_{X,R,A} \\ &= \langle \hat{X}(t - t_0^+) \rangle_{X,R,A} + \frac{2^{-1}}{1 - 2^{-1} e^{-\lambda_X^- T_G}} e^{-\lambda_X^-(t-t_0)} \langle \hat{X}(T_G) \rangle_{X,R,A}\end{aligned}\quad (\text{S41})$$

$$\begin{aligned}\langle R(t) \rangle_{R,A} &= \langle \hat{R}(t - t_0^+) \rangle_{R,A} + \frac{1}{2} e^{-\lambda_R^-(t-t_0)} \langle R(t_0^-) \rangle_{R,A} \\ &= \langle \hat{R}(t - t_0^+) \rangle_{R,A} + \frac{2^{-1}}{1 - 2^{-1} e^{-\lambda_R^- T_G}} e^{-\lambda_R^-(t-t_0)} \langle \hat{R}(T_G) \rangle_{R,A}.\end{aligned}\quad (\text{S42})$$

Here $\hat{X}(t - t_0^+)$ represents the amount of protein resulting from synthesis in an actual generation, whereas $X(t_0^-)$ is the amount of protein in the mother cell just before cell division. The precise expressions for $\hat{X}(t - t_0^+)$ and $\hat{R}(t - t_0^+)$ are given by

$$\begin{aligned}\langle \hat{X}(t - t_0^+) \rangle_{X,R,A} &= \frac{\mathcal{A}}{\lambda_R^-} \left[\frac{1}{\lambda_X^-} \left(1 - e^{-\lambda_X^-(t-t_0^+)} \right) - \frac{1 - 2^{-1} Z_R^{(1)}}{\lambda_R^- - \lambda_X^-} \left(e^{-\lambda_X^-(t-t_0^+)} - e^{-\lambda_R^-(t-t_0^+)} \right) \right] \\ \langle \hat{R}(t - t_0^+) \rangle_{R,A} &= \langle A \rangle_{B_{eq}} \frac{\lambda_R^+}{\lambda_R^-} \left[1 - e^{-\lambda_R^-(t-t_0^+)} \right].\end{aligned}$$

We obtain the expressions for the mean amount of protein and mRNA given in Eqns. (13), (14) in the main text.

The **transcriptional contribution** is given by

$$\langle (\langle X(t) \rangle_X - \langle X(t) \rangle_{X,R})^2 \rangle_{R,A} \quad (\text{S43})$$

$$= \underbrace{\left\langle (\langle X(t_0^+) \rangle_X - \langle X(t_0^+) \rangle_{X,R})^2 \right\rangle_{R,A}}_{T_1} e^{-2\lambda_X^-(t-t_0)} \quad (\text{S44})$$

$$+ \underbrace{2 \int_{t_0}^t \langle \langle X(t_0^+) \rangle_X \lambda_X(t') - \langle X(t_0^+) \rangle_{X,R} \langle \lambda_X(t') \rangle_R \rangle_{R,A} e^{-\lambda_X^-(t-t_0)} e^{-\lambda_X^-(t-t')} dt'}_{T_2} \quad (\text{S45})$$

$$+ \underbrace{\int_{t_0}^t \int_{t_0}^t \langle \lambda_X(t') \lambda_X(t'') - \langle \lambda_X(t') \rangle_R \langle \lambda_X(t'') \rangle_R \rangle_{R,A} e^{-\lambda_X^-(t-t'')} e^{-\lambda_X^-(t-t')} dt' dt''}_{T_3}. \quad (\text{S46})$$

From the invariance $t \rightarrow t + T_G$ we get an expression for the first term

$$T_1(t - t_0) = \frac{1}{4} \frac{T_2(T_G) + T_3(T_G)}{1 - \frac{1}{4} e^{-2\lambda_X^- T_G}} e^{-2\lambda_X^-(t-t_0)}. \quad (\text{S47})$$

Correlations between actual and previous generations are defined by

$$T_2(t - t_0) = \langle A \rangle \frac{\lambda_R^+}{\lambda_R^-} (\lambda_X^+)^2 B_{eq} \frac{1}{2} \left[1 - \frac{1}{4} e^{-(\lambda_R^- + \lambda_X^-) T_G} \right]^{-1} \left\{ \frac{1}{\lambda_R^- + \lambda_X^-} \left[1 - e^{-(\lambda_R^- + \lambda_X^-) T_G} \right] \right. \quad (S48)$$

$$\left. - \frac{1}{\lambda_X^-} \left[1 - e^{-\lambda_X^- T_G} \right] e^{-\lambda_R^- T_G} (1 - 2^{-1} Z_R^{(1)}) \right\} \frac{1}{\lambda_R^- - \lambda_X^-} \left[e^{-2\lambda_X^- (t-t_0)} - e^{-(\lambda_R^- + \lambda_X^-)(t-t_0)} \right].$$

Noise from the new generation is given by

$$T_3(t - t_0) = 2Q_X B_{eq} \left\{ (1 - 2^{-1} Z_R^{(1)}) \left[-\frac{\lambda_X^+}{\lambda_R^- - 2\lambda_X^-} \left(e^{-2\lambda_X^- (t-t_0)} - e^{-\lambda_R^- (t-t_0)} \right) \right. \right. \quad (S49)$$

$$\left. + \frac{\lambda_X^+}{\lambda_R^- - \lambda_X^-} \left(e^{-2\lambda_X^- (t-t_0)} - e^{-(\lambda_X^- + \lambda_R^-)(t-t_0)} \right) \right] + \frac{\lambda_X^+}{2(\lambda_R^- + \lambda_X^-)} \left(1 - e^{-2\lambda_X^- (t-t_0)} \right)$$

$$\left. - \frac{\lambda_X^- \lambda_X^+}{(\lambda_R^-)^2 - (\lambda_X^-)^2} \left(e^{-2\lambda_X^- (t-t_0)} - e^{-(\lambda_X^- + \lambda_R^-)(t-t_0)} \right) \right\}.$$

The major steps of the calculations to obtain $T_2(t - t_0)$ and $T_3(t - t_0)$ are shown in the Appendix.

The partial **contribution** to gene expression noise arising **from promoter activation** is given by

$$\langle (\langle X(t) \rangle_{X,R} - \langle X(t) \rangle_{X,R,A})^2 \rangle_A \quad (S50)$$

$$= \underbrace{\langle \langle X(t_0^+) \rangle_{X,R}^2 - \langle X(t_0^+) \rangle_{X,R,A}^2 \rangle_A}_{J_1} e^{-2\lambda_X^- (t-t_0)} \quad (S51)$$

$$+ 2 \underbrace{\int_{t_0^+}^t \langle \langle X(t_0^+) \rangle_{X,R} \langle \lambda_X(t') \rangle_R - \langle X(t_0^+) \rangle_{X,R,A} \langle \lambda_X(t') \rangle_{R,A} \rangle_A}_{J_2} e^{-\lambda_X^- (t-t_0)} e^{-\lambda_X^- (t-t')} dt' \quad (S52)$$

$$+ \underbrace{\int_{t_0^+}^t \int_{t_0^+}^t \langle \langle \lambda_X(t') \rangle_R \langle \lambda_X(t'') \rangle_R - \langle \lambda_X(t') \rangle_{R,A} \langle \lambda_X(t'') \rangle_{R,A} \rangle_A}_{J_3} e^{-\lambda_X^- (t-t'')} e^{-\lambda_X^- (t-t')} dt' dt'' \quad (S53)$$

Term J_3 can be split up into

$$\begin{aligned}
J_3(t-t_0) &= \phi \int_{t_0^+}^t \int_{t_0^+}^t \int_{-\infty}^{t'} \int_{-\infty}^{t''} \left[e^{-\gamma_A |\tau' - \tau''|} \left(\frac{1}{2} \right)^{D(t', \tau')} e^{-\lambda_R^-(t' - \tau')} \left(\frac{1}{2} \right)^{D(t'', \tau'')} \right. \\
&\quad \left. \times e^{-\lambda_R^-(t'' - \tau'')} e^{-\lambda_X^-(t - t')} e^{-\lambda_X^-(t - t'')} \right] d\tau'' d\tau' dt'' dt' \\
&= 2\phi \left[\underbrace{\int_{t_0^+}^t \int_{t_0^+}^{t'} \int_{t_0^+}^{t'} \int_{-\infty}^{t_0^+} [\cdot] d\tau'' d\tau' dt'' dt'}_{I_1} + \underbrace{\int_{t_0^+}^t \int_{t_0^+}^{t'} \int_{-\infty}^{t_0^+} \int_{-\infty}^{t_0^+} [\cdot] d\tau'' d\tau' dt'' dt'}_{I_2} \right. \\
&\quad \left. + \underbrace{\int_{t_0^+}^t \int_{t_0^+}^{t'} \int_{-\infty}^{t_0^+} \int_{t_0^+}^{t''} [\cdot] d\tau'' d\tau' dt'' dt'}_{I_3} + \underbrace{\int_{t_0^+}^t \int_{t_0^+}^{t'} \int_{t_0^+}^{t'} \int_{t_0^+}^{t''} [\cdot] d\tau'' d\tau' dt'' dt'}_{I_4} \right] \quad (S54)
\end{aligned}$$

and equivalently **term** J_2 reads

$$\begin{aligned}
J_2(t-t_0) &= \phi \int_{t_0^+}^t \int_{-\infty}^{t_0^+} \int_{-\infty}^{t'} \int_{-\infty}^{t''} \left[e^{-\gamma_A |\tau' - \tau''|} \left(\frac{1}{2} \right)^{D(t', \tau')} e^{-\lambda_R^-(t' - \tau')} \left(\frac{1}{2} \right)^{D(t'', \tau'')} \right. \\
&\quad \left. \times e^{-\lambda_R^-(t'' - \tau'')} e^{-\lambda_X^-(t - t')} e^{-\lambda_X^-(t - t'')} \left(\frac{1}{2} \right)^{D(t_0^+, t'')} \right] d\tau'' d\tau' dt'' dt' \quad (S55)
\end{aligned}$$

with

$$\phi = \text{var}(A) (\lambda_X^+)^2 (\lambda_R^+)^2 B_{eq}^2 = \frac{\text{var}(A)}{\langle A \rangle^2} \mathcal{A}^2. \quad (S56)$$

The term $J_2(t-t_0)$ is calculated for the asymptotic cases of interest whereas $J_3(t-t_0)$ can be given by the formulas for $I_1 - I_4$. The explicit expressions to obtain the single terms for $I_1 - I_4$ are derived in an equivalent way to the transcriptional contribution (see Appendix).

Calculations lead to the explicit expression for **term** I_1 , i.e.

$$\begin{aligned}
I_1 &= \phi_1 \left\{ \frac{1}{2\lambda_X^- - (\gamma_A + \lambda_R^-)} \left[e^{-(\gamma_A + \lambda_R^-)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \right. \\
&\quad - \frac{1}{\lambda_X^- - \gamma_A} \left[e^{-(\gamma_A + \lambda_X^-)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \\
&\quad - \frac{1}{2(\lambda_X^- - \lambda_R^-)} \left[e^{-2\lambda_R^-(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \\
&\quad \left. + \frac{1}{\lambda_X^- - \lambda_R^-} \left[e^{-(\lambda_R^- + \lambda_X^-)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \right\} \quad (S57)
\end{aligned}$$

with ϕ_1 given by

$$\phi_1 = \frac{1}{2} \frac{1}{(\lambda_R^-)^2 - \gamma_A^2} Z_{RA} \frac{1}{\lambda_X^- - \lambda_R^-}. \quad (S58)$$

Term I_2 is given by

$$I_2 = \frac{1}{4} \left\{ \frac{1}{\lambda_R^- + \gamma_A} \frac{1}{\lambda_R^-} Z_R^{(2)} - \frac{1}{(\lambda_R^-)^2 - \gamma_A^2} \frac{e^{-(\lambda_R^- + \gamma_A)T_G} - e^{-2\lambda_R^- T_G}}{(1 - 2^{-1}e^{-(\gamma_A + \lambda_R^-)T_G})(1 - 2^{-2}e^{-2\lambda_R^- T_G})} \right\} \\ \times \frac{1}{2(\lambda_X^- - \lambda_R^-)^2} \left[e^{-\lambda_R^-(t-t_0)} - e^{-\lambda_X^-(t-t_0)} \right]^2. \quad (\text{S59})$$

Term I_3 is calculated to be of the form

$$I_3 = \phi_3 \left\{ \frac{1}{\lambda_X^- - \gamma_A} \frac{1}{2\lambda_X^- - (\gamma_A + \lambda_R^-)} \left[e^{-(\lambda_R^- + \gamma_A)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \right. \\ \left. + \left(\frac{1}{\lambda_X^- - \lambda_R^-} \left[\frac{1}{\lambda_X^- - \lambda_R^-} - \frac{1}{\lambda_X^- - \gamma_A} \right] \right) \left[e^{-(\lambda_R^- + \lambda_X^-)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \right. \\ \left. - \frac{1}{2(\lambda_X^- - \lambda_R^-)^2} \left[e^{-2\lambda_R^-(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \right\} \quad (\text{S60})$$

with

$$\phi_3 = \frac{1}{2} \frac{1}{(\lambda_R^-)^2 - \gamma_A^2} Z_{RA}. \quad (\text{S61})$$

The explicit formula for **term** I_4 is given by

$$I_4 = \int_{t_0^+}^t \int_{t_0^+}^{t'} \int_{t_0^+}^{t''} \int_{t_0^+}^{t'''} [\cdot] d\tau'' d\tau' dt'' dt' \\ = \underbrace{\int_{t_0^+}^t \int_{t_0^+}^{t'} \int_{t_0^+}^{t''} \int_{t_0^+}^{t'''} [\cdot] d\tau'' d\tau' dt'' dt'}_{H_1} + \underbrace{\int_{t_0^+}^t \int_{t_0^+}^{t'} \int_{t''}^{t'} \int_{t_0^+}^{t'''} [\cdot] d\tau'' d\tau' dt'' dt'}_{H_2} \quad (\text{S62})$$

with

$$H_1 = \frac{2}{\lambda_R^- + \gamma_A} \\ \times \left\{ \frac{1}{2\lambda_R^-(\lambda_X^- + \lambda_R^-)} \left(\frac{1}{2\lambda_X^-} \left[1 - e^{-2\lambda_X^-(t-t_0)} \right] - \frac{1}{\lambda_X^- - \lambda_R^-} \left[e^{-(\lambda_R^- + \lambda_X^-)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \right) \right. \\ \left. + \left(\frac{1}{2\lambda_R^-(\lambda_X^- - \lambda_R^-)} + \frac{1}{(\lambda_R^- - \gamma_A)(\lambda_X^- - \gamma_A)} - \frac{1}{\lambda_R^- - \gamma_A} \frac{1}{\lambda_X^- - \lambda_R^-} \right) \right. \\ \times \frac{1}{\lambda_X^- - \lambda_R^-} \left[e^{-(\lambda_R^- + \lambda_X^-)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \\ \left. - \left(\frac{1}{2\lambda_R^-(\lambda_X^- - \lambda_R^-)} - \frac{1}{\lambda_R^- - \gamma_A} \frac{1}{\lambda_X^- - \lambda_R^-} \right) \frac{1}{2(\lambda_X^- - \lambda_R^-)} \left[e^{-2\lambda_R^-(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \right. \\ \left. - \frac{1}{(\lambda_R^- - \gamma_A)(\lambda_X^- - \gamma_A)} \frac{1}{2\lambda_X^- - (\gamma_A + \lambda_R^-)} \left[e^{-(\gamma_A + \lambda_R^-)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \right\} \quad (\text{S63})$$

and

$$\begin{aligned}
H_2 = & \frac{1}{\lambda_R^- + \gamma_A} \frac{1}{\lambda_R^- - \gamma_A} \\
& \times \left\{ \left(\frac{1}{\lambda_X^- + \gamma_A} - \frac{1}{\lambda_X^- + \lambda_R^-} \right) \frac{1}{2\lambda_X^-} \left[1 - e^{-2\lambda_X^-(t-t_0)} \right] \right. \\
& + \left(\frac{1}{\lambda_X^- - \lambda_R^-} - \frac{1}{\lambda_X^- + \gamma_A} \right) \frac{1}{\lambda_X^- - \gamma_A} \left[e^{-(\gamma_A + \lambda_X^-)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \\
& - \left(\frac{1}{\lambda_X^- - \lambda_R^-} - \frac{1}{\lambda_X^- - \gamma_A} \right) \frac{1}{2\lambda_X^- - (\lambda_R^- + \gamma_A)} \left[e^{-(\lambda_R^- + \gamma_A)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \\
& \left. + \left(\frac{1}{\lambda_X^- + \lambda_R^-} - \frac{1}{\lambda_X^- - \gamma_A} \right) \frac{1}{\lambda_X^- - \lambda_R^-} \left[e^{-(\lambda_R^- + \lambda_X^-)(t-t_0)} - e^{-2\lambda_X^-(t-t_0)} \right] \right\}. \quad (S64)
\end{aligned}$$

III. ASYMPTOTIC ANALYSIS OF THE PARTIAL CONTRIBUTIONS

A. Case I: Short mRNA lifetime, long protein lifetime $(\lambda_R^-)^{-1} \ll T_G \ll (\lambda_X^-)^{-1}$

This is the most physiological case and we obtain the asymptotic expressions for the mean molecule number within on cell cycle given in Eqns. (15), (16) in the main text. The **transcriptional contribution** can be split up into the correlations between the actual and previous generations (T_2) and the noise from the new generation (T_3) and we obtain for the asymptotic case $(\lambda_X^-)^{-1} \gg T_G$

$$T_2(t - t_0) = \frac{1}{2} \mathcal{A} \frac{\lambda_X^+}{(\lambda_R^-)^3} \left(1 - e^{-\lambda_R^-(t-t_0)} \right). \quad (S65)$$

and

$$T_3(t - t_0) = 2\mathcal{A} \frac{\lambda_X^+}{(\lambda_R^-)^2} \left[(t - t_0) - \frac{3}{2} \frac{1}{\lambda_R^-} \left(1 - e^{-\lambda_R^-(t-t_0)} \right) \right]. \quad (S66)$$

Including the contribution from T_1 and assuming $t - t_0 \gg (\lambda_R^-)^{-1}$, we obtain for the transcriptional contribution to gene expression noise the expression of Eqn. (19) in the main text

The single expression for the **contributions from promoter activation** are given by

$$\begin{aligned}
J_2(t - t_0) = & \mathcal{A}^2 \frac{\text{var}(A)}{\langle A \rangle^2} \frac{1}{(\lambda_R^-)^2 - (\gamma_A)^2} \left(\frac{1}{\gamma_A^2} Z_A^{(1)} \left[1 - e^{-\gamma_A(t-t_0)} \right] \right. \\
& \left. - \frac{\gamma_A}{(\lambda_R^-)^3} \left[1 - e^{-\lambda_R^-(t-t_0)} \right] \right) \quad (S67)
\end{aligned}$$

and

$$J_3(t - t_0) = 2\phi H_2 = 2\mathcal{A}^2 \frac{\text{var}(A)}{\langle A \rangle^2} \frac{1}{(\lambda_R^-)^2 - (\gamma_A)^2} \left[\frac{1}{\gamma_A}(t - t_0) - \frac{1}{\gamma_A^2}(1 - e^{-\gamma_A(t-t_0)}) \right] \quad (\text{S68})$$

Assuming $t - t_0 \gg (\gamma_A)^{-1}, (\lambda_R^-)^{-1}$ and including the contributions from J_1 , we obtain the noise contribution from promoter activation given in Eqn. (20) in the main text.

B. Case II: Long mRNA and protein lifetime $(\lambda_R^-)^{-1}, (\lambda_X^-)^{-1} \gg T_G$

The terms for the **transcriptional contribution** read

$$T_2(t - t_0) = \frac{1}{3} \mathcal{A} \lambda_X^+ T_G^2 (t - t_0). \quad (\text{S69})$$

and noise from the new generation is given by

$$T_3(t - t_0) = \frac{1}{3} \mathcal{A} \lambda_X^+ (t - t_0)^3. \quad (\text{S70})$$

The **contributions from promoter activation** read

$$J_2(t - t_0) = \mathcal{A}^2 \frac{\text{var}(A)}{\langle A \rangle^2} \frac{Z_A^{(1)}}{\gamma_A} \left\{ 8T_G^2(t - t_0) + \frac{4}{3\gamma_A} T_G(t - t_0) - \frac{1}{\gamma_A^2} (1 - e^{-\gamma_A(t-t_0)}) \frac{1 - e^{-\gamma_A T_G}}{1 - \frac{1}{4} e^{-\gamma_A T_G}} \right\} \quad (\text{S71})$$

and

$$J_3(t - t_0) = 2\mathcal{A}^2 \frac{\text{var}(A)}{\langle A \rangle^2} (I_1 + I_2 + I_3 + I_4) \quad (\text{S72})$$

with

$$I_1 = \frac{1}{4\gamma_A^2} Z_A^{(1)} (t - t_0)^2 \quad (\text{S73})$$

$$I_2 = \frac{1}{6\gamma_A} \left[2T_G - \frac{Z_A^{(1)}}{\gamma_A} \right] (t - t_0)^2 \quad (\text{S74})$$

$$I_3 = \frac{Z_A^{(1)}}{2\gamma_A^2} \left\{ \frac{1}{\gamma_A} (1 - e^{-\gamma_A(t-t_0)}) - \frac{t - t_0}{\gamma_A} + \frac{(t - t_0)^2}{2} \right\} \quad (\text{S75})$$

$$I_4 = \frac{(t - t_0)^3}{3\gamma_A} - \frac{(t - t_0)^2}{2\gamma_A^2}. \quad (\text{S76})$$

In the asymptotic limit (for $T_G, (t - t_0) \gg \gamma_A^{-1}$) we obtain Eqns. (23) and (24) in the main text for the noise due to transcription and promoter activation, respectively.

C. Case III: Short mRNA and short protein lifetime $(\lambda_R^-)^{-1}, (\lambda_X^-)^{-1} \ll T_G$

The mean amounts of protein and mRNA are given by Eqns. (25) and (26) in the main text. The terms for the **transcriptional contribution** are given in the limit $T_G \gg (\lambda_R^-)^{-1}, (\lambda_X^-)^{-1}$ by

$$T_2(t - t_0) = \frac{2}{3} \mathcal{A} \frac{\lambda_X^+}{\lambda_R^-((\lambda_R^-)^2 - (\lambda_X^-)^2)} \left(e^{-2\lambda_X^-(t-t_0)} - e^{-(\lambda_R^- + \lambda_X^-)(t-t_0)} \right). \quad (\text{S77})$$

and

$$\begin{aligned} T_3(t - t_0) = Q_X B_{eq} \left\{ \left[-\frac{\lambda_X^+}{\lambda_R^- - 2\lambda_X^-} \left(e^{-2\lambda_X^-(t-t_0)} - e^{-\lambda_R^-(t-t_0)} \right) \right. \right. \\ + \frac{\lambda_X^+}{\lambda_R^- - \lambda_X^-} \left(e^{-2\lambda_X^-(t-t_0)} - e^{-(\lambda_X^- + \lambda_R^-)(t-t_0)} \right) \left. \right] + \frac{\lambda_X^+}{2(\lambda_R^- + \lambda_X^-)} \left(1 - e^{-2\lambda_X^-(t-t_0)} \right) \\ \left. - \frac{\lambda_X^- \lambda_X^+}{(\lambda_R^-)^2 - (\lambda_X^-)^2} \left(e^{-2\lambda_X^-(t-t_0)} - e^{-(\lambda_X^- + \lambda_R^-)(t-t_0)} \right) \right\}. \quad (\text{S78}) \end{aligned}$$

The only **contribution from promoter activation** derives from the I_4 term ($I_1 = I_2 = I_3 = 0$). As the fluctuations decay very fast we also have no contribution from previous generations ($J_1 = 0$) and no correlations between the previous and actual generation ($J_2 = 0$).

We obtain

$$J_3(t - t_0) = 2\mathcal{A}^2 \frac{\text{var}(A)}{\langle A \rangle^2} I_4 = \frac{\mathcal{A}^2 \text{var}(A)}{\langle A \rangle^2 \lambda_R^- \lambda_X^- (\gamma_A + \lambda_R^-)(\lambda_R^- + \lambda_X^-)} \left(1 + \frac{\lambda_R^-}{\gamma_A + \lambda_X^-} \right). \quad (\text{S79})$$

Eqns. (27) and (28) of the main text are obtained in the limit $t - t_0 \gg (\gamma_A)^{-1}$.

IV. SIMULATING PARTIAL NOISE CONTRIBUTIONS

We investigate the partial contributions (terms of Eqn. (S33)) to gene expression noise by the following simulation protocol. If we assume a constant activator $B_{eq} = \text{const}$, there are no contributions from activator binding. In the more general case, i.e. if $B(t) \neq \text{const}$, it can be easily included in the simulations.

As mentioned before, we can also neglect the noise contribution from translation. The remaining terms (transcription) and (promoter activation) only differ in their averaging.

In order to calculate the partial contributions to gene expression noise, a correct averaging of the simulation results is necessary:

Let g, m be the number of simulations of promoter-*on*/*off* trajectories and mRNA trajectories, respectively. Each protein trajectory has a unique promoter-*on*/*off* and mRNA trajectory, so p_{ij} is the protein trajectory resulting from the i th promoter-*on*/*off* and the j th mRNA trajectory and can be identified with $\langle X(t) \rangle_X$.

Averaging over all mRNA trajectories belonging to one promoter-*on*/*off* trajectory, we obtain the mean protein trajectory belonging to the i th promoter-*on*/*off* realization, i.e. $\langle X(t) \rangle_{X,R} = \frac{1}{m} \sum_{j=1}^m p_{ij}$.

Averaging over all mRNA trajectories which were obtained from all promoter-*on*/*off* trajectories we have $\langle X(t) \rangle_{X,R,A} = \frac{1}{g} \sum_{i=1}^g \frac{1}{m} \sum_{j=1}^m p_{ij}$.

We obtain for the second moment of transcription (σ_R^2) and promoter activation (σ_A^2) of gene expression noise on the protein level:

$$\begin{aligned} \sigma_R^2 &= \langle (\langle X(t) \rangle_X - \langle X(t) \rangle_{X,R})^2 \rangle_{R,A,B} = \langle (\langle X(t) \rangle_X - \langle X(t) \rangle_{X,R})^2 \rangle_{R,A} \\ &= \frac{1}{g} \sum_{i=1}^g \frac{1}{m} \sum_{j=1}^m (p_{ij} - \frac{1}{m} \sum_{j=1}^m p_{ij})^2 \end{aligned} \quad (S80)$$

$$\begin{aligned} \sigma_A^2 &= \langle (\langle X(t) \rangle_{X,R} - \langle X(t) \rangle_{X,R,A})^2 \rangle_{A,B} = \langle (\langle X(t) \rangle_{X,R} - \langle X(t) \rangle_{X,R,A})^2 \rangle_A \\ &= \frac{1}{g} \sum_{i=1}^g (\frac{1}{m} \sum_{j=1}^m p_{ij} - \frac{1}{g} \sum_{i=1}^g \frac{1}{m} \sum_{j=1}^m p_{ij})^2. \end{aligned} \quad (S81)$$

To illustrate that this averaging method mimics the correct formulas, if mRNA and protein degradations are large compared to the generation time (*case III*), i.e. the system equilibrate very fast after cell division, Table S1 summarizes the mean amount of mRNA and protein and the noise contributions from transcription and promoter activation (averaged over the last half of the cell cycle). If, however, one of these two rates is small compared to the generation time (*case I* and *II*), the simulation/averaging routine becomes much more complex and very time consuming. To assure that the analytical expressions are correct, we calculated some values of the integral representation of the partial contributions (Eqns. (S52), (S53)) with Mathematica and compared them to the explicit formulas (Eqns. (S67),(S71)) and (S57)-(S64)). For *case I* and *II*, the precise formulas evaluated at three different time points coincide with the integral representations at these time points.

V. APPENDIX

We derive the explicit formulas for the transcriptional contribution, i.e. Eqns. (S45,S46) and give some calculations leading to (S48) and (S49). The contribution from promoter activation is obtained in an equivalent way.

$$\begin{aligned}
T_2(t - t_0) &= (\lambda_X^+)^2 \int_{t_0}^t \int_{-\infty}^{t_0^-} \underbrace{\langle \langle R(t'')R(t') \rangle_R - \langle R(t'') \rangle_R \langle R(t') \rangle_R \rangle_A}_{\langle R(t'') \rangle_R e^{-\lambda_R^-(t'-t'')} 2^{-D(t',t'')}} \rangle_A \\
&\quad \times e^{-\lambda_X^-(2t-t'-t'')} 2^{-D(t_0^-,t'')} dt'' dt' \\
&= \frac{(\lambda_X^+)^2 \lambda_R^+}{\lambda_R^-} \langle A \rangle B_{eq} \int_{t_0^-}^t \int_{-\infty}^{t_0^-} \left[1 - e^{-\lambda_R^-(t''-t_0)} (1 - 2^{-1} Z_R^{(1)}) \right] \\
&\quad \times e^{-(\lambda_X^- + \lambda_R^-)(t'-t'')} e^{-2\lambda_X^-(t-t')} \frac{1}{2} 4^{-D(t_0^-,t'')} dt'' dt' \\
&= \frac{(\lambda_X^+)^2 \lambda_R^+}{2\lambda_R^-} \langle A \rangle B_{eq} \int_{t_0^-}^t \sum_{k=0}^{\infty} 4^{-k} \int_{t_0^-(k+1)T_G}^{t_0-kT_G} \left[1 - e^{-\lambda_R^-(t''-(t_0-(k+1)T_G))} (1 - 2^{-1} Z_R^{(1)}) \right] \\
&\quad \times e^{-(\lambda_X^- + \lambda_R^-)(t'-t'')} e^{-2\lambda_X^-(t-t')} dt'' dt' \\
&= \frac{(\lambda_X^+)^2 \lambda_R^+}{2\lambda_R^-} \langle A \rangle B_{eq} \int_{t_0^-}^t \left[\frac{1 - e^{-(\lambda_X^- + \lambda_R^-)G}}{(\lambda_X^- + \lambda_R^-)(1 - \frac{1}{4} e^{-(\lambda_X^- + \lambda_R^-)T_G})} \right. \\
&\quad \left. - \frac{e^{-\lambda_R^- T_G} - e^{-(\lambda_X^- + \lambda_R^-)T_G}}{\lambda_X^-(1 - \frac{1}{4} e^{-(\lambda_X^- + \lambda_R^-)T_G})} (1 - 2^{-1} Z_R^{(1)}) \right] * e^{-(\lambda_X^- + \lambda_R^-)t'} e^{2\lambda_X^-(t-t')} dt' \\
&= (S48)
\end{aligned}$$

$$\begin{aligned}
T_3(t - t_0) &= 2(\lambda_X^+)^2 \int_{t_0}^t \int_{t_0}^{t'} \langle R(t'') \rangle_{R,A} e^{-\lambda_R^-(t'-t'')} e^{-\lambda_X^-(2t-t'-t'')} dt'' dt' \\
&= \frac{2(\lambda_X^+)^2 \lambda_R^+}{\lambda_R^-} \langle A \rangle B_{eq} \int_{t_0}^t \int_{t_0}^{t'} \left[(2^{-1} Z_R^{(1)} - 1) e^{-\lambda_R^-(t''-t_0)} + 1 \right] e^{-\lambda_R^-(t'-t'')} \\
&\quad \times e^{-\lambda_X^-(2t-t'-t'')} dt'' dt' \\
&= 2\lambda_X^+ Q_X B_{eq} \int_{t_0}^t \left[(2^{-1} Z_R^{(1)} - 1) e^{-\lambda_R^-(t'-t_0)} (1 - e^{-\lambda_X^-(t'-t_0)}) e^{-2\lambda_X^-(t-t')} \right. \\
&\quad \left. + \frac{\lambda_X^-}{\lambda_R^- + \lambda_X^-} (1 - e^{-(\lambda_X^- + \lambda_R^-)(t'-t_0)}) e^{-2\lambda_X^-(t-t')} \right] dt' \\
&= 2\lambda_X^+ Q_X B_{eq} \left\{ \frac{1}{2(\lambda_X^- + \lambda_R^-)} + \frac{1 - 2^{-1} Z_R^{(1)}}{\lambda_R^- - 2\lambda_X^-} e^{-\lambda_R^-(t-t_0)} \right. \\
&\quad + e^{-2\lambda_R^-(t-t_0)} \left[(1 - 2^{-1} Z_R^{(1)}) \left(\frac{1}{\lambda_R^- - 2\lambda_X^-} - \frac{1}{\lambda_X^- + \lambda_R^-} \right) - \frac{1}{2(\lambda_X^- + \lambda_R^-)} \right. \\
&\quad \left. \left. - \frac{\lambda_X^-}{(\lambda_X^-)^2 - (\lambda_R^-)^2} \right] + e^{-(\lambda_R^- + \lambda_X^-)(t-t_0)} \left(2^{-1} Z_R^{(1)} - 1 \right) \frac{1}{\lambda_X^- + \lambda_R^-} + \frac{\lambda_X^-}{(\lambda_X^-)^2 - (\lambda_R^-)^2} \right\} \\
&= (S49)
\end{aligned}$$

	analytics	simulations
$\langle R(t) \rangle_{\mathbf{I}}$	5.33	5.31 ± 0.03
$\langle X(t) \rangle_{\mathbf{I}}$	106.67	106.37 ± 0.39
η_R^2	0.1071	$0.1074 \pm 4.6 \times 10^{-4}$
η_A^2	0.2402	$0.2443 \pm 2.7 \times 10^{-3}$

TABLE S1: Comparison of the analytical and simulation results of *case III* with the kinetic parameters $\lambda_A^+ = 0.2$, $\lambda_A^- = 0.3$, $\lambda_R^+ = 2$, $\lambda_R^- = 0.15$, $\lambda_X^+ = 4$, $\lambda_X^- = 0.2$ and $B_{eq} = 1$. The cells divide after 100 minutes.

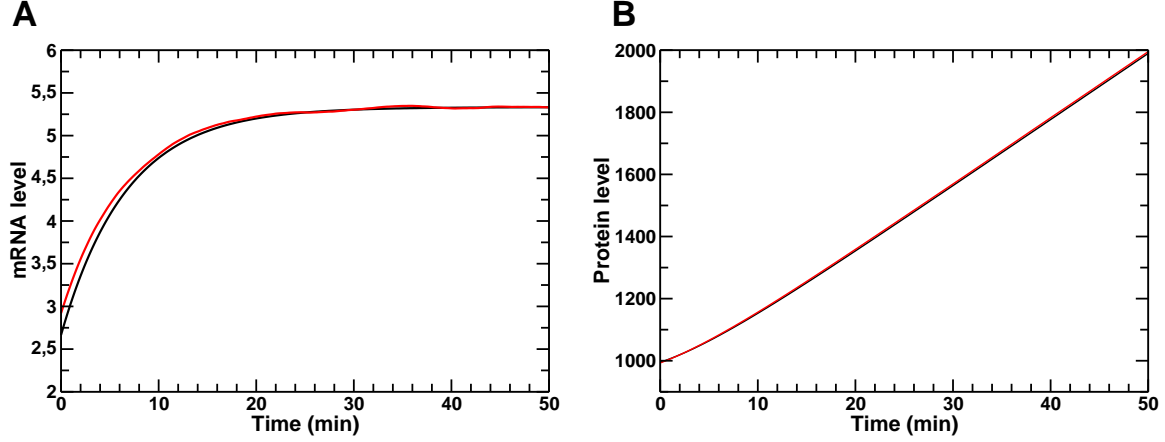


FIG. S1: Comparison of analytical results (black lines) with simulations (red lines). (A) mean amount of mRNA and (B) mean amount of protein. The kinetic parameters are chosen from *case I*, when mRNA lifetime is short and protein lifetime is long compared to the generation time, i.e. $\lambda_A^+ = 0.2$, $\lambda_A^- = 0.3$, $\lambda_R^+ = 2$, $\lambda_R^- = 0.15$, $\lambda_X^+ = 4$ and $\lambda_X^- = 1 \times 10^{-5}$. The activator is constantly bound, i.e. $B_{eq} = 1$, but the same results are also obtained if $B_{eq} = q < 1$ (simulations not shown). We fix the cell cycle to $t_0 = 0$, $T = 50$ (min) and sampling frequency to $dt = 0.2$. To obtain a good agreement between simulations and analytics, we simulate and average over 100 cells, each with 200 promoter-on/-off trajectories. For each promoter-on/-off trajectory we generate 200 mRNA trajectories with the corresponding protein trajectories.